Imaging and Mitigation of Travelling Ionospheric Disturbances

Bolmgren, Karl

Award date:
2020

Awarding institution:
University of Bath

Imaging and Mitigation of Travelling Ionospheric Disturbances

submitted by

Karl Henrik Anton Bolmgren

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Electronic & Electrical Engineering

September 2020

Summary

Travelling ionospheric disturbances, or TIDs, are wavelike features propagating in the ionosphere. TIDs are studied for many reasons, such as their effects on GNSS navigation and their connection to natural disasters like earthquakes and tsunamis. Sensitive instruments and reliable techniques are required to accurately image and detect the ionospheric perturbations. The primary purpose of this thesis is to evaluate and advance the capabilities of tomographic methods to image TIDs for scientific purposes, which also has a potential application to GNSS positioning in the presence of TIDs. This thesis therefore also quantifies the potential impact of TIDs on state-of-the-art ionospheric correction services, for e.g. Network Real-Time Kinematic (N-RTK) positioning, that occur when the corrections are interpolated for an approximate rover position. Building on these results, a novel TID-mitigation strategy is also tested. Simulations with different methods of interpolation for N-RTK ionospheric corrections show that TIDs can induce errors large enough to merit attention. For positioning in the presence of TIDs, a weighted least-squares interpolation technique is modified to adapt to estimated TID directions and wavelengths. The new methods are shown, in simulation tests and a case study, to decrease the number of large interpolation errors that may impede fast integer ambiguity resolution. Electron density maps generated by ionospheric tomography can also provide an alternative approach to TID mitigation. The MIDAS (Multi-Instrument Data Analysis System) tomography algorithm is therefore tested to determine its suitability for TID imaging. As an initial test, a case study is presented a where MIDAS is used to image a large-scale TID over North America occurring during a geomagnetic storm. The resulting images are verified with in-orbit measurements and ionosonde soundings. Further tests are conducted with simulated data generated with varying TIDs characteristics and satellite geometries, including geostationary orbits. The results show that most TIDs are reconstructed well by MIDAS, with the smaller MSTIDs a possible exception. Electron density images generated from these results can be used for TID mitigation in positioning, and also aid studies into TID and gravity wave generation mechanisms. Together, the results from studies emphasise the importance of satellite and receiver geometry in TID observation and mitigation. The simulation-based tomography results show the benefits of including geostationary satellite geometry in ionospheric tomography, while results from the LSTID case study and interpolation simulations illustrate the importance of a well distributed ground receiver network.

Contents

Introduction

State of the art navigation is today provided by several Global Navigation Satellite Systems (GNSS), whose key element - the satellites - orbit the Earth around twenty thousand kilometres away. GNSS have, since the inception of GPS and GLONASS in the 1980s, become increasingly ubiquitous to activities on Earth. Highly accurate and precise positioning, navigation and timing made available by modern GNSS techniques and used in different industries are required to be not only exact, but also reliable and fast in different situations, environments and conditions. The inherent variability and unpredictability of disturbances in the ionosphere make ionospheric studies central to these endeavours.

The most prevalent perturbations in the ionosphere are Travelling Ionospheric Disturbances (TIDs), and these are of interest both for their effects on positioning and for our understanding of the interconnections between the surface of the earth, the atmosphere and space. TIDs are observable as wave-like signatures in the ionospheric electron density and can be initiated by processes in the Earth's magnetosphere as well as in the neutral atmosphere below.

Activity below the ionosphere can interface with ionospheric plasma through Atmospheric Gravity Waves (AGWs), buoyancy waves in the neutral atmosphere. If an AGW reaches the ionosphere it can set the ionospheric ions in motion and be detectable as a TID (Hines, 1960). Such waves are relevant not only as a potential issue for high accuracy GNSS, but also as a mechanism of energy- and momentum transfer between different layers of the atmosphere. This is of particular interest for developers of modern climate models, for which the correct inclusion of AGWs is important (e.g. Garcia et al., 2017).

While some AGWs begin in the troposphere by weather systems, sharp temperature gradients or interactions between winds and topography (e.g. Nappo, 2013, and references therein), they can also be initiated under more dramatic circumstances. Recently the potential to use TIDs as a proxy to detect acoustic-gravity waves originating from

Tsunamis (Kherani et al., 2016; Savastano et al., 2017) and earthquakes (Kherani et al., 2016; Yang et al., 2017) have been explored in order to develop new early warning systems. Other events like volcano eruptions (Shults et al., 2016), rocket launches (Noble, 1990) and nuclear explosions (Hines, 1967; Breitling et al., 1967) have also been known to cause similar perturbations in the ionosphere, and the generation mechanisms of TIDs on Earth and in space are widely studied in their own right.

Other TIDs have their origin above the neutral atmosphere. Geomagnetic storms, results of space weather, cause large currents in the polar regions of the Earth. These currents are hypothesised to drive the generation of large TIDs that propagate from the poles toward the equator (Borries et al., 2017). Space weather, which is intimately connected to solar activity, is not the only way in which the Sun causes TIDs on Earth; the heat gradients created by the passage of the solar terminator also generates local AGWs and TIDs(Hernández-Pajares, Juan and Sanz, 2006).

Advances in our knowledge of topics such as the ionosphere and space weather can benefit GNSS technology by enabling the development of new ionospheric models. In return, developments in GNSS has been a consistent driver and enabler of space- and atmospheric research. Because the received GNSS signals are affected by atmospheric delays, receivers with a known position can be used to recover information about the medium the signal has passed through. Typically, state of the art real-time high accuracy GNSS positioning algorithms use networks of reference receivers. These provide the necessary corrections used to position a roving receiver quickly and accurately. The large geographic areas covered by these networks provide opportunities to study the ionosphere and troposphere. This research lies at this intersection, where the monitoring, imaging and detection of ionospheric disturbances meet the mitigation of their effects on high-accuracy GNSS positioning.

The potential errors arising from TIDs in real-time kinematic positioning algorithms are investigated in Chapter 5. In order to improve the application of ionospheric corrections in the presence of TIDs, a modified interpolation technique is also tested and compared to existing methods. While effective, some errors remain, especially for sparser networks. An alternative to interpolation of ionospheric corrections is considered in the subsequent chapters 6 and 7: *ionospheric tomography*.

The development of denser receiver networks has made imaging tools like tomography increasingly powerful for detecting TIDs. Tomography, as an ionospheric monitoring method, is described in Chapter 3. Chapters 6 and 7 implement and evaluate ionospheric GNSS tomography on varying scales of TIDs in different circumstances.

Chapters 5 – 7 are collected in part II, which contains the main research of the thesis. Some concepts that are essential to what is discussed in part II are introduced in part I. This part includes the fundamentals of GNSS, the ionosphere, tomography and the current state of TID monitoring. The final part of the thesis, part III, concludes the analysis in a discussion of the methods and results. The focus of part III lies in what this means for the future of TID-studies with – and for the benefit of – GNSS.

List of publications

Publications included in the thesis

Tomographic Imaging of Traveling Ionospheric Disturbances using GNSS and Geostationary Satellite Observations, **K. Bolmgren**, C. Mitchell, J. Bruno, G. Bust. JGR: Space Physics, Feb 2020. doi: 10.1029/2019JA027551

Tomographic Imaging of a Large Scale TID during the Halloween Storm of 2003, **K. Bolmgren**, C. Mitchell, T. Pinto Jayawardena, G. Bust, J. Bruno. Annales Geophysicae, in press. doi: 10.5194/angeo-2020-26

Additional publications not included in the thesis

A realistic simulation framework to evaluate ionospheric tomography, J. Bruno, C. Mitchell, **K. Bolmgren**, B. Witvliet. Advances in Space Research, Feb 2020. doi: 10.1016/j.asr.2019.11.015

GPS loss of lock statistics over Brazil during the 24th solar cycle, J. Damaceno*, **K. Bolmgren***, J. Bruno*, G. De Franceschi, C. Mitchell, M. Cafaro. (*equal authorship), Advances in Space Research, Jul 2020. doi: 10.1016/j.asr.2020.03.041

Quality analysis of dual-frequency smartphone-based ionospheric TEC measurements: what can be achieved? J. Bruno, F. Darugna, **K. Bolmgren**, J. Wübbena, M. Schmitz, C. Mitchell. Annals of Geophysics, Oct 2020. doi: 10.4401/ag-8517

Mitigation of Severe Weather Events and TID Impact on the Interpolation of SSR Atmospheric Parameters, F. Darugna, **K. Bolmgren**, M. Schmitz, S. Schön, J. Wübbena, G. Wübbena, J. Bruno, C. Mitchell. Advances in Space Research, in review.

Acknowledgements

Firstly, I would like to thank my supervisor Prof Cathryn Mitchell for her invaluable support, guidance and inspiration during these three years of challenging and interesting work, which enabled my development as a researcher and a person.

This work was funded by the Horizon 2020 Marie Skłodowska-Curie Actions grant agreement No. 722023, TREASURE, and I am grateful for the opportunities made available to me through this programme.

I would also like to thank the people at Spirent Communications in Paignton, MIT Haystack Laboratory, NASA's Goddard Space Flight Center and Geo++ for their hospitality during my secondments and research visits.

My time as a doctoral student would not have been so enjoyable without my colleagues and friends at the University of Bath and the TREASURE network. Special mentions go out to Jon Bruno, who shared this journey with me from start to finish, and to Dr Talini Pinto Jayawardena, whose heroic reviewing efforts saved this text from illegibility.

Lastly, I would like to thank my family for their encouragement and support.

Karl Bolmgren
15 September 2020

Part I

1 The ionosphere

The ionosphere can be found in the upper part of the atmosphere where neutral molecules are ionised to form a plasma. It is located between approximately 80 km and 1000 km in altitude above the surface of the Earth. This means that the overwhelming majority of the ionosphere is above the Kármán line (100 km in altitude) in what is classified as outer space, and the study of the ionosphere is part of the field of space physics.

Although the ionosphere has been visible since pre-historical times in the form of the aurora borealis and the aurora australis, scientific study of the ionosphere took off with the advent of wireless radio technology. The existence of the ionosphere was evidenced by Marconi's famous transmission over the Atlantic Ocean in 1901, which was made possible by the radio waves being reflected off of the ionosphere. However, this was only proven a quarter of a decade later in 1924, independently by Appleton and Barnett in the UK, and Breit and Tuve in the USA. Another major leap in the field of space physics came with the beginning of the space age, when humankind for the first time could make in situ measurements of the properties of space using rockets and satellites. In fact, the earliest rocket capable of reaching space, the German V2, was used as a sounding rocket as early as March 1944 to measure atmospheric pressure, ultraviolet radiation and take air samples (Seibert, 2007).

This chapter aims to give a short introduction to the fundamental concepts and general characteristics of the ionosphere, as well as the effects on radio wave transmissions in the ionosphere. The textbooks by Hargreaves (1992) and Prölss (2004) were consulted during the preparation of this chapter.

1.1 Structure and formation of the ionosphere

The ions and free electrons in the ionospheric plasma are largely created by ionisation in the ionosphere. There is therefore an approximately equal number of ions and free

electrons in a given volume, so that the ion and electron densities fulfil the relation

$$N_i \approx N_e. \tag{1}$$

This means that the ionospheric plasma is quasi-neutral, and that the electric fields from the positive and negative charges of the ions and electrons cancel out, making the plasma neutrally charged on average. The spatial scale where this approximation holds is related to the *Debye length*. This is the maximum charge separation between inhomogeneities arising from thermal movement of electrons with temperature T_e. With the Debye length, λ_D, can be written

$$\lambda_D = \sqrt{\frac{\epsilon_0 k_B T_e}{q_e^2 N_e}} \tag{2}$$

where ϵ_0 is the permittivity of free space, k_B is the Boltzmann constant and q_e is the elementary charge. In the ionosphere, λ_D varies from a few mm in the F region to around a cm in the D region.

There is a large part of the atmosphere that is neutral also at ionospheric heights and these neutral atoms and molecules serve as the source of the charged electrons and ions. This generally happens through *photoionisation*, where an amount of extreme ultraviolet (EUV) or X-Ray radiation hf from the Sun detaches an electron e^- from the neutral atom X:

$$X + hf \rightarrow e^- + X^+. \tag{3}$$

EUV radiation is more intense at higher altitudes, where very little of it has been absorbed. The neutral atmosphere decreases in density with height, so there are fewer neutral atoms available to be ionised as the altitude increases. In contrast, at lower altitudes where the neutral density is higher, less radiation is available due to absorption by the atmosphere it has already passed through. This creates a layer of maximum ionisation at some height where there is a balance between ionising radiation and density of neutral molecules in the atmosphere. Since there are several different species of neutral gases in the atmosphere, and the Sun emits radiation in a spectrum, the ionosphere can consist of up to four different such layers. These are the D, E, F1 and F2 layers illustrated by the ionospheric profiles in Figure 1. This basic structure of the ionosphere was first described by Chapman (1931), and the mathematical formulation of an individual layer following his work is called a *Chapman layer*.

Given that solar radiation is the main source of ionisation, after the Sun sets the rate of photoionisation drops and most of the ions are lost due to *recombination*. Recombination

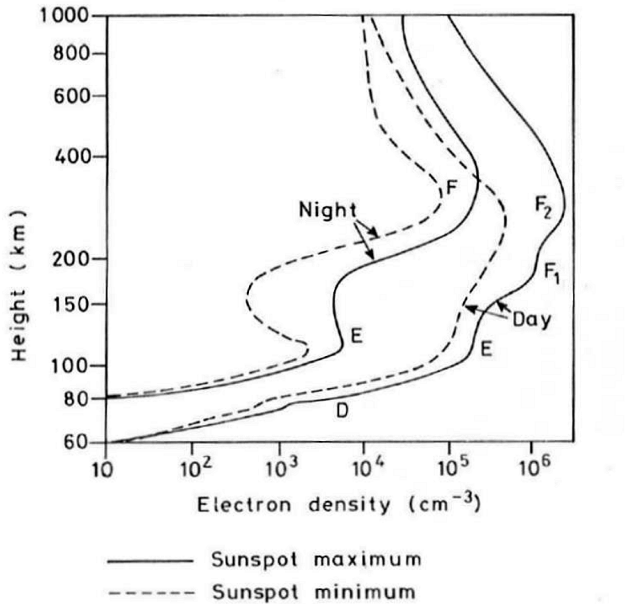

Figure 1: Mid-latitude ionospheric electron density variations by height. Solid lines correspond to solar maximum, and hatched lines solar minimum. Reproduced from the textbook by Hargreaves (1992).

is a reaction between a positively charged ion and a negatively charged free electron that produces a neutral atom or molecule. The most important type of recombination is dissociative recombination, which happens in two stages as a positive ion X^+ first interacts with a neutral molecule Y_2 before recombining with an electron:

$$X^+ + Y_2 \rightarrow YX^+ + Y \tag{4}$$

$$YX^+ + e^- \rightarrow Y + X \tag{5}$$

The rate of the process described in (4) is dependent on both the availability of ions X^+ and neutrals Y_2. At high altitudes neutral molecules become scarcer, and in the F2 layer this is the limiting reaction that determines the rate of dissociative recombination. At the F1 layer, in contrast, the neutral density is such that the process in (5) is the limiting reaction. With the approximation that the electron and ion densities are equal, the rate of (5) depends linearly on the electron density and (4) on the squared electron density. Recombination is therefore slower in the F2 layer, and it typically stays ionised throughout the night when the F1 peak loses its ionisation due to recombination.

As the level of photoionisation of the dayside ionosphere is dependent on the amount of ionising radiation from the Sun, it follows the *solar cycle*. The solar cycle is an

approximately 11-year periodic cycle in solar activity, and during the peak of a cycle the levels of ionisation in the ionosphere are higher than during low solar activity, as illustrated in Figure 1. A common indicator of solar activity is the number of *sunspots*, which appear as visually dark spots on the Sun and are indicative of increased magnetic flux on the Sun. During periods of high solar activity, it is more likely that activity from *solar flares* occurring on the Sun cause increased ionisation in the ionosphere termed a *sudden ionospheric disturbance*. Another phenomenon linked with solar flares are *Coronal Mass Ejections* (CMEs), which are expulsions of plasma from the Sun's own atmosphere, or *corona*. If a CME hits the Earth, it can have significant effects on the Earth's *magnetosphere*, ionosphere and current systems, resulting in what are known as *geomagnetic storms*.

The Earth's magnetic field can be roughly approximated as a dipole with a moment slightly off-axis from the Earth's axis of rotation. While the magnetic field lines at mid- and low-latitude are closed, the magnetic field lines near the magnetic poles are open and coupled with the *interplanetary magnetic field* of the *solar wind*. The solar wind is the ever present and variable stream of plasma emanating from the Sun in all directions, and it carries with it magnetic field lines originating from the Sun. The solar wind plasma consists of electrons, protons and alpha particles, but the ratio between alpha particles and protons remains uncertain (Borovsky, 2020). These energetic particles can travel from the solar wind and enter the Earth's atmosphere along coupled open field lines of the geomagnetic field and become an important source of ionisation in the polar atmosphere. Photons emitted from interactions between these energetic particles and the neutral molecules of the atmosphere are responsible for the aurorae visible in the polar regions.

1.2 Travelling ionospheric disturbances

Travelling Ionospheric Disturbances (TIDs) are wave-like perturbations in the local electron density of the ionosphere. They originate from buoyancy waves in the neutral atmosphere called Atmospheric Gravity Waves (AGWs), or simply gravity waves. In this work, these terms are used interchangeably. This section includes a basic description of gravity waves, and how they relate to TIDs.

It is well known that sound propagates in the atmosphere. Sound waves, however, are only a special case of the more general acoustic-gravity waves present in the atmosphere. They can be understood as the result of a restoring buoyancy force acting upon a displacement of gas in the atmosphere. The initial displacement can be caused by

external factors such as particle precipitation, and internal processes such as weather systems causing pressure and temperature gradients. An example of a gravity wave at low altitude is shown in Figure 2, where the low temperature nodes of the wave are made visible by condensation in the cold nodes of the wave.

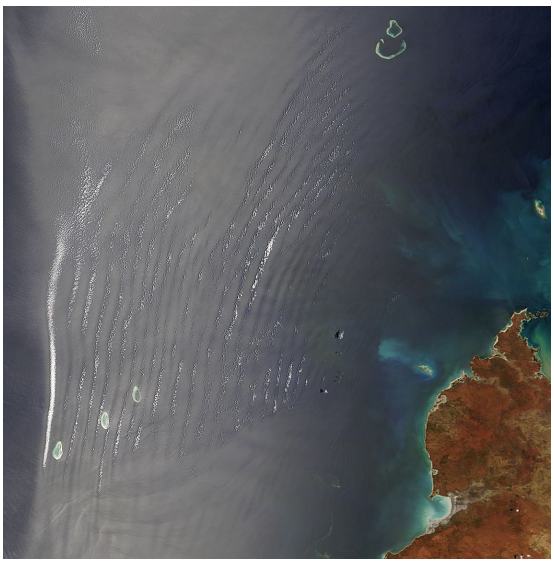

Figure 2: Atmospheric gravity wave made visible by condensation. Image credit: NASA/GSFC/MODIS Land Rapid Response Team and Jeff Schmaltz `https://www.nasa.gov/multimedia/imagegallery/image_feature_484.html`

The dispersion relation for acoustic-gravity waves was first described by Hines (1960) as

$$\omega^4 - \omega^2 s^2 k^2 + \omega_b^2 s^2 k_h^2 - \omega^2 \omega_a^2 = 0 \tag{6}$$

where s is the sound velocity, ω is the angular frequency, k is the wave vector, and k_h is the horizontal component of the wave vector. The acoustic cut-off angular frequency ω_a and the isothermal *Brunt-Väisälä angular frequency* (or *buoyancy frequency*) ω_b can be written

$$\omega_a = \gamma g / 2s \tag{7}$$

$$\omega_b = (\gamma - 1)^{\frac{1}{2}} g / s, \tag{8}$$

where γ is the heat capacity ratio of the air and g is the acceleration due to gravity. The purely acoustic case is found when the effect of gravity is neglected so that $\omega \gg \omega_a > \omega_b$ and (6) simplifies to

$$k = \frac{\omega}{s}. \tag{9}$$

When instead $\omega \ll \omega_b$, the effect of compressibility is neglected, and the dispersion

relation simplifies to

$$k_z^2 = \frac{\omega_b^2}{\omega^2} k_h^2 - \frac{\omega_a^2}{s^2} \tag{10}$$

which describes a pure gravity wave. Waves with angular frequency $\omega < \omega_b$ are sometimes called *internal gravity waves*.

As gravity waves reach ionospheric heights, motions in the neutral atmosphere interact with the ionosphere. The most significant effect of this interaction is that of neutral particles conveying their motion to ions by collision. Since the ions are charged their movements will follow the geomagnetic field lines. The resulting perturbations manifest as waves in the ionosphere, i.e. TIDs. The TID perturbation in electron density can range from tenths of a percent to several percent, and AGWs have periods of minutes to hours, and wavelengths from tens to thousands of km. TIDs are typically classified as either Large Scale (LSTID), Medium Scale (MSTID) or Small Scale (SSTID) TIDs. While LSTIDs and MSTIDs are caused by AGWs, SSTIDs are linked with acoustic waves and are outside the scope of the work in this thesis. Typical values for different classifications of TIDs can be found in Table 1. More information on the research into TIDs can be found in the reviews by Yeh and Liu (1974), Hunsucker (1982) and Hocke and Schlegel (1996).

Nomenclature	Velocity $v_{\varphi,h}$ (m/s)	Period T (min)	Wavelength λ_h (km)
LSTID	400–1000	30–180	$\geq$1000
MSTID	100–250	15–60	100–1000
SSTID	300–3000	2–5	

Table 1: Typical horizontal phase velocities ($v_{\varphi,h}$), periods (T) and horizontal wavelengths (λ_h) of TIDs, adapted from the review by Hunsucker (1982).

1.3 Ionospheric radio propagation

An important effect of the free electrons in the ionosphere is that they interact with propagating electromagnetic waves. The propagation of electromagnetic waves is characterised by the phase (v_φ) and group velocity (v_g) defined by the following relations

$$v_\varphi = \frac{\omega}{k} \tag{11}$$

and

$$v_g = \frac{\partial \omega}{\partial k} \tag{12}$$

where $\omega = 2\pi f$ is the angular wave frequency and $k = 2\pi/\lambda$ is the wave number. When travelling in a medium, these velocities are governed by the respective indices of

refraction,

$$n_\varphi = \frac{c}{v_\varphi}, \quad n_g = \frac{c}{v_g}, \tag{13}$$

where c is the speed of light in vacuum and n_φ and n_g are the phase and group refractive indices, respectively. These indices can be found using the Appleton-Hartree Equation 14, which gives the phase refraction-index of electromagnetic waves in a magnetised cold plasma,

$$n_\varphi^2 = 1 - \frac{\omega_p^2/\omega^2}{1 - i\frac{\nu}{\omega} - \frac{\frac{\omega_g^2}{2\omega^2}\sin^2\theta}{1-\frac{\omega_p^2}{\omega^2}-i\frac{\nu}{\omega}} \pm \frac{1}{1-\frac{\omega_p^2}{\omega^2}-i\frac{\nu}{\omega}}\left(\frac{\omega_g^4}{4\omega^4}\sin^4\theta + \frac{\omega_g^2}{\omega^2}\cos^2\theta(1-\frac{\omega_p^2}{\omega^2}-i\frac{\nu}{\omega})^2\right)^{1/2}} \tag{14}$$

and the relation (15) between group and phase refractive indices, derived from the definitions (11), (12) and (13)

$$n_g = \frac{\partial(\omega n_\varphi)}{\partial\omega}. \tag{15}$$

In Equation 14, ν is the electron collision frequency; θ is the angle between the background magnetic field and the propagation direction of the wave; i is the imaginary unit;

$$\omega_g = \frac{q_e B}{m_e} \tag{16}$$

is the electron *gyrofrequency* characterising the rotation of an electron (with mass m_e and charge q_e) in a magnetic field with magnitude B;

$$\omega_p = \sqrt{\frac{N_e q_e^2}{\epsilon_0 m_e}} \tag{17}$$

is the electron *plasma frequency*, the resonant frequency of a plasma with electron density N_e; and ϵ_0 is the permittivity of free space.

The somewhat unwieldy expression given in (14) can be simplified to give a better understanding of ionospheric radio propagation if the effects of absorption and magnetic field are disregarded ($\nu = 0$, $B = 0$), reducing it to

$$n_\varphi^2 = 1 - \frac{\omega_p^2}{\omega^2} \tag{18}$$

which, when ω_p^2/ω^2 is close to zero, can be approximated as

$$n_\varphi = \sqrt{1 - \frac{\omega_p^2}{\omega^2}} \approx 1 - \frac{\omega_p^2}{2\omega^2} = 1 - \frac{\kappa}{f^2} N_e, \tag{19}$$

where

$$\kappa \triangleq \frac{q_e^2}{8\pi^2 \epsilon_0 m_e} \approx 40.3\,[\mathrm{Hz}^2\mathrm{m}^3]. \tag{20}$$

The approximation (19) holds for L-band (1-2 GHz) signals propagating the ionosphere, where typically $\omega_p^2/\omega^2 \approx 10^{-2}$.

The index of refraction for the group velocity, from equations (15) and (19) now becomes

$$n_g = 1 + \frac{\kappa}{f^2} N_e\,. \tag{21}$$

It is evident from these equations that the ionospheric index of refraction depends on the frequency of the incoming wave. Media with this property are called *dispersive*.

The difference in sign between equations (19) and (21) means that while the phase velocity of an electromagnetic wave is increased in the ionosphere, the group velocity of the wave is decreased. To illustrate this effect, one can consider the time τ it takes for a radio signal to propagate through the ionosphere along a ray-path $\mathcal{R}$ with length ρ:

$$\tau = \int_{\mathcal{R}} \frac{dl}{v} = \int_{\mathcal{R}} \frac{n}{c} dl, \tag{22}$$

substituting the refractive index, n, with equations (19) and (21) for a given signal, leaves

$$\tau = \int_{\mathcal{R}} \frac{1}{c} \left(1 \pm \frac{\kappa}{f^2} N_e\right) dl. \tag{23}$$

After integrating along the signal path, the propagation time of the signal through the ionosphere becomes

$$\tau = \frac{1}{c} \left(\rho \pm \frac{\kappa}{f^2} \int_{\mathcal{R}} N_e\, dl\right). \tag{24}$$

If the apparent distance were to be estimated using the time measurement as $\rho_{est} = c\tau$, the ionosphere would add an additional term $\pm\frac{\kappa}{f^2}\int N_e dl$. The sign of this term depends on whether the distance is calculated using the group or phase measurements. The factor $\int N_e dl$ is called the *Total Electron Content* (TEC), and is typically given in TEC units (1 TECu = 10^{16} electrons per m^2).

Just like light in the visible spectrum, when an electromagnetic wave in the radio-

frequency domain reaches a sharp boundary between media with different refractive indices, it can be either refracted or reflected. For a HF-radio signal broadcast from Earth, the gradual increase in ionospheric electron density with altitude results in a similar effect. The ray is gradually changing direction as a result of refraction, as it propagates further up in the ionosphere, until it has reversed direction and exits the ionosphere back towards Earth or passes through the ionosphere. Reflection can happen when the used frequency is below the maximum plasma frequency (or *critical frequency*) of the ionosphere. Signals in the L-band, however, are far above the typical ionospheric plasma frequency and experience minimal ray bending.

1.4 Summary

In this chapter, the ionosphere has been introduced as the ionised part of the Earth's atmosphere. The ionosphere is created primarily through photoionisation by radiation and particle precipitation form the Sun. It is intimately connected to the geomagnetic field and solar wind, which connects eruptions on the Sun to geomagnetic storms and other disturbances in the ionosphere.

The ionosphere can also be perturbed by sources in the neutral atmosphere. TIDs, a common type of ionospheric disturbance, are caused by atmospheric gravity waves reaching the ionospheric. They can be observed as propagating wave-like perturbations in the electron density and span a wide range of spatial and temporal scales.

The ionosphere has a profound impact on electromagnetic waves by slowing down the group velocity, increasing the phase velocity and bending ray-paths as they propagate the region. Since the ionosphere is dispersive, these effects depend on the frequency of the incident wave as well as the ionospheric electron density.

2 | Global navigation satellite systems

Global Navigation Satellite Systems (GNSS) are systems of orbiting satellite constellations with accompanying ground segments that are designed to provide an accurate global positioning service. The most widely known GNSS is undoubtedly the *Global Positioning System*, or GPS. It was the first of several satellite-based navigation systems like the *GLObal NAvigation Satellite System* (GLONASS) and the more recent *BeiDou navigation satellite System* (BDS) and *Galileo* system. These GNSS share many general characteristics, and this chapter aims to outline the basics of GNSS in terms of navigation and ionospheric monitoring. The textbook by Teunissen and Montenbruck (2017) was consulted in preparing this chapter.

2.1 GNSS segments

Each GNSS has a *space segment*, a *control segment*, and a *user segment*. The core of the space segment of a given GNSS is a constellation of satellites in Medium Earth Orbit (MEO), at approximately 20,000 km in altitude, designed to enable at least four satellites to be in view of the user anywhere on Earth. For GPS, Galileo and BDS, these orbits have approximate inclinations of 55° with respect to the equatorial plane. GLONASS has a different inclination, 64.8°, which results in higher satellite elevations at high latitudes. The GPS space segment currently consists of 31 available in-orbit satellites, while GLONASS has 24 and Galileo 22 active navigation satellites. In addition to 27 operational MEO satellites, BDS has 6 satellites in Geostationary Earth Orbit (GEO) and 10 in Inclined Geosynchronous Orbit (IGSO).

A satellite broadcasts signals on two to four separate frequencies within the L-band, carrying Pseudo-Random Noise (PRN) codes, unique to the satellite and making it possible to identify the satellite from which the signal is broadcast and its time of

transmission.

Apart from the PRN, the signals carry additional information needed for basic positioning, including clock corrections, the approximate orbit and basic atmospheric parameters. The PRN is generated by high quality on-board oscillators, which are core components of any GNSS. The other parameters are calculated in the control segment using observations from its global monitoring stations and uploaded to the satellites from a central monitoring station.

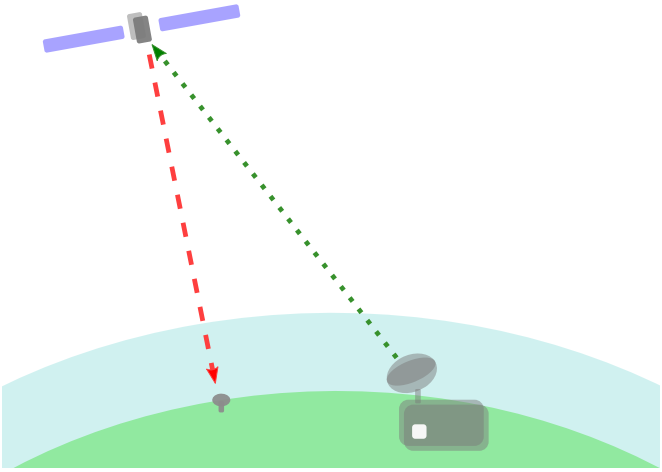

Figure 3: Illustration of the standard information flow between parts of the space, control and user segments of GNSS. The ground segment collects and computes the corrections, which are uploaded (green, dotted line) to the GNSS satellites, whose broadcasts (red, dashed line) are picked up by GNSS receivers of the user segment.

Lastly, the user segment consists of GNSS receivers used for positioning and navigation. These vary in quality, from the multi-frequency high-quality receivers with calibrated antennas used in research and professional surveying, to low-cost single frequency chips found in mass-market products. The user segment receives data from the MEO GNSS satellites, as illustrated in Figure 3, but can also obtain supporting information from Satellite-Based Augmentation Systems (SBAS) satellites in Geosynchronous Orbits (GSO), Ground-Based Augmentation Systems (GBAS), or via an internet connection. This information typically includes detailed or regional corrections that are not suitable for transmission from the global MEO satellites. Examples of advanced positioning techniques that rely on accurate supporting information are Precise Point Positioning (PPP) and Real-Time Kinematic (RTK) positioning. The basics of RTK is covered in section 2.4. The user receiver in these cases are commonly called the roving receiver, or *rover*, to distinguish it from reference stations receivers used to gather corrections.

2.2 Basic positioning

The basic idea of GNSS is to use estimates of the range from satellites to receiver to determine the receiver position. It is, however, not the distance that is measured, but the travel time of the signal between satellite and receiver. The apparent range P can then be inferred from the travel time τ by

$$P = c\tau \tag{25}$$

where c is the speed of light. In order to measure the signal delay τ, the PRN code of the GNSS signal used. This is a repeating binary sequence that is encoded onto the signal, and which is unique to each satellite. A receiver replicates these codes, uses auto-correlation of the replicated code and the received code to determine τ in Equation 25.

The unknown receiver position has three degrees of freedom in space, so three known ranges and satellite positions can in theory be used to determine the receiver position by *trilateration*, illustrated in Figure 4. However, the time measurements used to obtain the ranges rely on highly accurate clocks in both receiver and satellite, and while GNSS satellites have high accuracy atomic clocks, receivers generally do not. To compensate, a fourth satellite measurement is needed to determine the extra degree of freedom arising from the unknown receiver clock error.

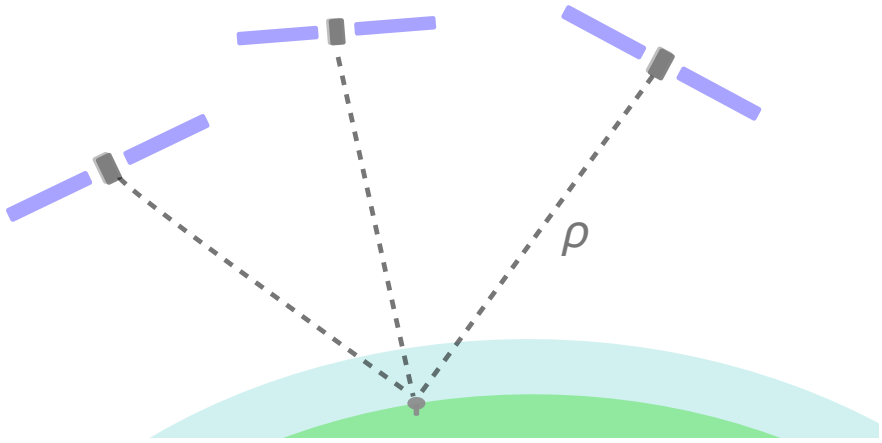

Figure 4: Illustration of the geometric range ρ between three satellites and one receiver. Three known distances can be used to determine the location of the receiver by trilateration.

In reality, the distance measured this way is not the true geometric distance, due to

many other errors, the most important of which are introduced in section 2.3. Therefore, P is known as the *pseudorange* to separate it from the geometrical range, ρ. These errors can, to varying degrees, be modelled or treated as part of receiver or satellite biases. Modelled solutions for these biases can be uploaded to satellites by the control segment and broadcast together with the satellite position as a navigation message, which is modulated on the carrier wave of the satellite signal.

2.3 GNSS observables

The primary measurement in GNSS is the pseudorange, which can be derived from code observations. In order to relate the pseudorange (P) and the geometric distance (ρ), clock, orbit, atmospheric, and other errors in the received signal must be accounted for:

$$P = \rho + c(dt_r - dt_s + \delta t_{rel}) + c(d_r - d_s) + I + T + \varepsilon. \tag{26}$$

Equation 26 is called the pseudorange *observation equation*, and contains the terms for receiver and satellite clock errors dt_r, dt_s, the corrections due to special and general relativity δt_{rel}, the instrument delays d_r, d_s, the tropospheric delay T, the combined errors due to multipath and receiver noise ε and the ionospheric delay

$$I = \frac{\kappa}{f^2}\text{TEC (see section 1.3).} \tag{27}$$

In basic single-frequency code-based positioning, parameters used to determine the TEC, as well as dt_s and T are usually taken from the navigation message while dt_r is estimated alongside the position from the pseudorange measurements.

Another important observable is the *carrier-phase*, φ, of the transmitted signal, which has a similar observation equation:

$$\varphi = \rho + c(dt_r - dt_s + \delta t_{rel}) + c(d_{r,\varphi} - d_{s,\varphi}) - I + T + \lambda N + \varepsilon_\varphi \tag{28}$$

where the additional term λN represents the unknown number of full carrier phase cycles at the start of phase tracking. N is known as the *integer ambiguity*. If the integer ambiguity is resolved, positioning derived from carrier phase measurements have a much higher accuracy than code-based positioning. Note that the ionospheric term I is negative in the carrier-phase observation equation. This is because the carrier-phase is not delayed but advanced in the ionosphere (as discussed in section 1.3).

While positioning is possible using single-frequency code or carrier-phase measurements

directly, observables are often combined in different linear combinations to suit specific applications. One such linear combination is called the *geometry-free* combination. It can be used to estimate ionospheric TEC and is defined as the difference between carrier-phase observations of two different frequencies:

$$\varphi_{1,2} \triangleq \varphi_1 - \varphi_2 \,, \tag{29}$$

which by (28) expands to the expression

$$\varphi_{1,2} = c(d_{r,1} - d_{r,2}) - c(d_{s,1} - d_{s,2}) - I_1 + I_2 + (\lambda_1 N_1 - \lambda_2 N_2) + \varepsilon_{1,2} \tag{30}$$

with the notation $\varepsilon_{1,2} \triangleq \varepsilon_1 - \varepsilon_2$. If the biases and ambiguities are resolved, (30) simplifies to

$$\varphi_{1,2} = I_2 - I_1 + \varepsilon_{1,2} \,. \tag{31}$$

By using the expression in (27) for the phase advances, TEC can then be estimated from the relation

$$\varphi_{1,2} = \kappa \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) \mathrm{TEC} + \varepsilon_{1,2} \,. \tag{32}$$

The integer ambiguities can be solved using, for example, the Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method (Teunissen, 1994). The inter-frequency biases, on the other hand, can be solved for using a least-squares approach and ionospheric models or images (e.g. Dear and Mitchell, 2006; Bruno et al., 2020).

While absolute TEC is needed for positioning corrections, the change in TEC with time is often of interest for ionospheric studies. If this is the parameter of interest, it is possible to largely remove the biases and ambiguities in (30) without integer ambiguity resolution by taking the difference between consecutive observations, under the assumption that the biases remain time-independent:

$$\Delta\varphi_{1,2} = \varphi_{1,2}(t_2) - \varphi_{1,2}(t_1) = \kappa \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) \Delta\mathrm{TEC} + \Delta\varepsilon_{1,2} \tag{33}$$

2.4 Real-time kinematic positioning

Real-Time Kinematic (RTK) positioning is the use of carrier phase measurements following the resolution of the integer ambiguities to achieve real time high accuracy (mm to cm) positioning relative to a reference station. Standard RTK uses *double-differenced* observables, which are pseudorange and carrier phase observations that have

been differenced relative to a reference satellite and relative to a reference receiver station:

$$P_i^{dd} = P_i - P_i^{ref} - (P_0 - P_0^{ref}) \tag{34}$$

$$\varphi_i^{dd} = \varphi_i - \varphi_i^{ref} - (\varphi_0 - \varphi_0^{ref}) \tag{35}$$

where the subscript $i = 0$ indicates the satellite selected as reference, and the superscript ref indicates a measurement at the reference receiver station. This procedure eliminates most of the additional terms in the observation equations (26) and (28), provided the receiver station is close enough to approximate the atmospheric delays I and T at the rover with the delays at the reference location. This leaves the integer ambiguity $N\lambda$ and stochastic error ε as the additional terms, and it is the successful resolution of integer ambiguities that results in the high accuracy of RTK.

In order to successfully resolve integer ambiguities, however, the distance between the rover and the reference station is required to be short, ideally less than 10-20 km. The problem can be solved at larger distances from the reference station by waiting for changing satellite geometry, which provides more independent observations. This takes time, however, and is therefore impractical for many applications.

A faster method to resolve ambiguities for roving receivers further away from the reference station is to use separate estimations of the ionospheric and tropospheric delays and satellite biases to speed up the ambiguity resolution. To make such estimations over a select region, a network of reference stations with GNSS receivers is needed. This approach is known as Network-RTK (N-RTK). N-RTK reference networks are typically required to have reference stations every 100-200 km to effectively provide corrections to receivers in the region covered by the network. As will be seen in Chapter 5, this makes N-RTK vulnerable to local changes in the ionosphere and troposphere, as the values estimated by the network need to be interpolated to the approximate position of the roving receiver before they can be applied to estimate the exact position.

2.5 Summary

This chapter describes the basic concept of GNSS - global satellite systems designed for positioning and navigation. Biases from many different sources need to be accounted for when employing GNSS for positioning. Importantly for this thesis, GNSS signals experience signal delay and phase advance in the ionosphere, which need to be accounted for to enable accurate positioning. However, the same effects on the signal can also

be used for ionospheric monitoring by retrieving TEC from the geometry-free linear combination of carrier phase observations on two separate frequencies.

High accuracy GNSS positioning can be attained in real time by using RTK or N-RTK, among other techniques. The methods covered here use the carrier phase and integer ambiguity resolution to accurately determine a position by using one or more reference stations at known locations.

3 | Ionospheric tomography

Computerised tomography, first introduced for the ionosphere by Austen et al. (1988), is a mathematical procedure used to image the interiors of objects using non-invasive measurements. This is called solving the *inverse problem*, defined in opposition to the *forward problem*, which determines the outcome of a process, e.g. light being diffracted by a lens, knowing the properties of the medium (lens) and initial conditions (angle of incidence). In the inverse problem, the outcome (in our example, the angle of diffraction) is known by measurements but the properties of the medium are to some degree unknown. Solving the inverse problem in tomography means imaging the interior properties of a medium by way of non-invasive measurements.

As described in section 1.3, the ionospheric medium has a profound effect on radio waves that propagate through it. The group delay and phase advance on trans-ionospheric signals are related to the TEC, and by measuring the TEC it is possible to estimate the ionospheric electron density on which these effects depend. Section 2.3 showed how the geometry-free combination of GNSS carrier phase measurements can be used to estimate the ionospheric TEC. More generally, a network of GNSS receivers with enough simultaneous observations can be used to estimate the ionospheric electron density distribution over a region and time period. This is usually termed Computerised Ionospheric Tomography (CIT), or ionospheric Radio Tomography (RT). Examples of tomography in other fields of research include seismic wave tomography, X-ray tomography and magnetic resonance imaging.

While it is now common practice to use GNSS observations in CIT, the technique was originally developed for TEC measurements from Low Earth Orbiting (LEO) ionospheric monitoring and navigation satellite systems like the American *TRANSIT* and Russian *CICADA* systems. Other data sources, such as GNSS occultation (e.g. Rius et al., 1997), airglow images (e.g. Kamalabadi et al., 2002) and ionosonde observations (e.g. Chartier et al., 2012) and has also since been used as inputs to CIT. A historical review of CIT

and ionospheric imaging can be found in the article by Bust and Mitchell (2008).

3.1 Solving the inverse problem

Ionospheric tomography takes integrated electron density measurements, TEC, as input and uses these to estimate the distribution of the ionospheric electron density. Solving the inverse problem in this case essentially means determining the distribution through which the TEC has been integrated, by only knowing the geometry of the ray and the TEC values.

In order to do this computationally the geometry of the problem needs to be discretised. A matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ represents the m ray-paths between receivers and satellites, discretised in n latitude-longitude-altitude voxels through which the rays connecting satellites and receivers pass. Each element a_{ij} of $\mathbf{A}$ is the length of the ith ray within the jth voxel. This matrix contains the geometry of the problem and is sometimes called the *kernel matrix* of the inverse problem. The TEC measurements associated with each ray-path are stored in a data vector $\mathbf{b} \in \mathbb{R}^m$, which is related to the discretised electron density $\mathbf{x} \in \mathbb{R}^n$ by

$$\mathbf{Ax} = \mathbf{b}. \tag{36}$$

This relation represents the discrete integration of the electron density along the ray-paths described by $\mathbf{A}$.

Solving Equation 36 for the electron density $\mathbf{x}$ means solving the inverse problem. In this case this gives an image of electron density based on the information in the TEC measurements $\mathbf{b}$, with the resolution depending on the number of voxels in $\mathbf{A}$. A simplified CIT geometry is shown in Figure 5.

Equation 36 can be written as a set of m equations to account for each ray-path

$$\sum_{j=1}^{n} a_{ij} x_j = b_i. \tag{37}$$

When $m \neq n$, i.e. there are unequal amounts of equations and unknowns, the equations cannot be solved for x_j analytically. A common approach instead is to select the solution which minimises the square of the residual,

$$f(\mathbf{x}, \mathbf{b}) = ||\mathbf{Ax} - \mathbf{b}||^2, \tag{38}$$

where f is called the *cost function*. This is the least-squares approach, and it is useful

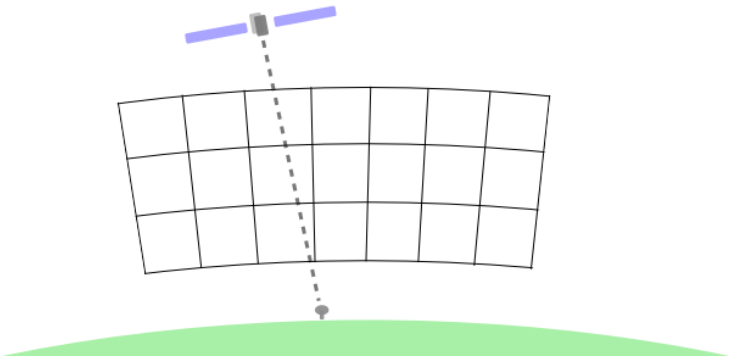

Figure 5: Simplified illustration of the geometry in a tomographic problem. The ionosphere is represented by latitude-longitude-altitude voxels through which passes the rays connecting satellites and receivers, each with an associated TEC value.

when there are more measurements than there are unknowns, i.e. $m < n$, and the problem is *overdetermined*. *Underdetermined* problems can instead be solved by selecting a cost function that minimises the norm,

$$f(\mathbf{x}) = ||\mathbf{x}||. \tag{39}$$

The minimum-norm solution is designed to minimise the complexity of the solution, rather than the misfit with regards to the data. Methods also exist that compromise between minimising the squared residual and the complexity. This is useful when problems are ill-conditioned, so that small errors in input can lead to large errors in output if the solution is not constrained, or *regularised*. Examples of such methods are *Tikhonov regularisation* (Tikhonov and Arsenin, 1977) and Truncated Singular Value Decomposition (TSVD) (Hansen, 1987).

As a useful notation, solutions to the inverse problem (36) can be written in the formalism of *generalised*, or *pseudoinverses*, $(\cdot)^{-\mathrm{g}}$,

$$\hat{\mathbf{x}} = \mathbf{A}^{-\mathrm{g}}\mathbf{b}. \tag{40}$$

where $\hat{\mathbf{x}}$ denotes the estimated solution. These are used because the regular inverse $(\cdot)^{-1}$ is not available for singular matrices $\mathbf{A}$. For example, the pseudoinverse following the least-squares approach can be found by minimising the cost function (38) with regards to $\mathbf{x}$, and becomes

$$\mathbf{A}^{-\mathrm{g}} = (\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal, \tag{41}$$

where $\mathbf{A}^\intercal$ is the transpose of $\mathbf{A}$.

3.1.1 Tikhonov regularisation

In the Tikhonov regularisation approach, the cost function is selected to combine the residual and model norm as

$$f(\mathbf{x}, \mathbf{b}) = ||\mathbf{A}\mathbf{x} - \mathbf{b}||^2 + \varrho||\mathbf{R}\mathbf{x}||^2, \tag{42}$$

where the choice of ϱ determines the trade-off between the relative importance of the model norm and residual terms, and $\mathbf{R}$ is selected according to the choice of regularisation. For example, if $\mathbf{R} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, this term becomes the regular norm used in the minimum-norm approach. Other common choices are the Jacobian $\mathbf{R} = \nabla$ and Hessian $\mathbf{R} = \nabla^2$ matrices, which if coupled with large ϱ force a smoother solution by minimising the first or second order gradients. By minimising $\mathbf{x}$ in (42), the Tikhonov regularisation pseudoinverse becomes

$$\mathbf{A}^{-\mathrm{g}} = (\mathbf{A}^\intercal\mathbf{A} + \varrho\mathbf{R}^\intercal\mathbf{R})^{-1}\mathbf{A}^\intercal. \tag{43}$$

3.1.2 The Moore-Penrose pseudoinverse and TSVD

The *Moore-Penrose pseudoinverse*, often designated by $(\cdot)^\dagger$, is calculated using Singular Value Decomposition (SVD), a generalisation of eigendecomposition for with non-square matrices, as

$$\mathbf{A}^\dagger = \mathbf{V}\mathbf{S}^{-\mathbf{1}}\mathbf{U}^\intercal. \tag{44}$$

The matrices on the right-hand side of (44) are given by the SVD of $\mathbf{A}$ and fulfil

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\intercal \tag{45}$$

where the square orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$ span the data and model spaces respectively, and the diagonal matrix $\mathbf{S}$ contains the singular values $s_i \geq 0$ of $\mathbf{A}$.

A stable solution can be achieved by truncating $\mathbf{S}$ so as to remove very small singular values, e.g. $s_i < s_1 \cdot 10^{-7}$. These singular values close to zero correspond to the model and data *null spaces*. Vectors in the data null space cannot be fitted to, and vectors in the model null space cannot be detected by the data. For example, a voxel pierced by no rays would fall into the model null space. Eliminating the data and model null spaces therefore results in a more well-posed inverse problem, and this constitutes the TSVD approach. In TSVD, the selection of the truncation threshold (in the example above

$s_1 \cdot 10^{-7}$) adjusts the regularisation akin to the choice of ϱ in Tikhonov regularisation.

3.2 Challenges in CIT

The ionosphere can be discretised in different ways, yielding different matrices $\mathbf{A}$ for the reconstruction. The simplest geometry is representing the ionosphere with a single thin shell that the electron density is mapped onto. When using this discretisation, the TEC of each ray is mapped to a single point on the ionospheric shell as vertical TEC (vTEC). To approximate the effect of the incident angle on how far the ray passes through the actual ionosphere, a *vTEC mapping function* is used. While the single thin shell approach leaves fewer unknowns as each pixel in the resulting estimate contains the vTEC of that location, electron density information along the altitude dimension is lost.

This limitation of the thin-shell approximation can be addressed using a more detailed representation of the ionosphere in three dimensions discretised as voxels, or multiple shells, instead of a single layer of pixels. This increases the number of unknowns in $\mathbf{x}$ without increasing the number of data points in $\mathbf{b}$, which is still equal to the number of receiver-satellite pairs. This results in inverse problems that are more underdetermined and may benefit from additional data.

An evenly distributed GNSS ground receiver network and rays that pierce the ionosphere more or less vertically give the data a good horizontal distribution, but a very poor vertical distribution. Measurements from low elevation angles suffer from multipath and propagate through a large part of the atmosphere, and thereby being subject to high noise levels are often discarded. To address this issue, three-dimensional ionospheric tomography methods sometimes use a-priori knowledge in the form of models to provide additional vertical structure. This information can be used, for example, as basis functions in a change of basis (see section 3.3, as an initial guess in an iterative algorithm. A modelled electron density distribution can also be used in a regularised solution by exchanging the minimum norm $||\mathbf{x}||$ for $||\mathbf{x} - \mathbf{x}_0||$, thereby tending the solution towards the modelled distribution $\mathbf{x}_0$ instead of the zero vector.

Another potential challenge for ionospheric tomography is the calibration of slant TEC observations. Instrumental biases (see section 2.3) are present in the phase observations, and these biases propagate to the TEC determined by the geometry-free combination. Equation 36 can be written to explicitly include these, as well as the undetermined integer ambiguities, as

$$\mathbf{A}\mathbf{x} = \mathbf{b} + \beta, \tag{46}$$

where the term

$$\beta = (c\Sigma d_{s,r} + \lambda_1 \mathbf{N}_1 - \lambda_2 \mathbf{N}_2) \left(\frac{\kappa}{f_2^2} - \frac{\kappa}{f_1^2} \right)^{-1} \tag{47}$$

incorporates the collected instrument biases $\Sigma d_{s,r}$ and integer ambiguity terms for each receiver-satellite link. It is possible to rewrite equation 46 in order to solve for this term as an additional unknown, although this leads to a more ill-conditioned inverse problem.

3.3 MIDAS

The Multi-Instrument Data Analysis System (MIDAS) is set of algorithms for tomography developed at the University of Bath that is capable of 2D (space) to 4D (time and 3D-space) tomographic imaging of the ionosphere and is also used for other media such as medical imaging. Here, the focus is on the ionospheric use. The MIDAS time-dependant inversion algorithm is described by Mitchell and Spencer (2003) and enables the use of differential inputs as an alternative to using bias-corrected measurements, or solving for biases as part of the inversion. The problem is instead formulated using the change in ray path geometry and the change in TEC recorded over a time window. This is done by taking the time difference (denoted by the Δ_t operator) of Equation 46 between consecutive time steps, giving

$$\Delta_t(\mathbf{A}\mathbf{x}) = \Delta_t(\mathbf{b} + \beta) \tag{48}$$

which expands to

$$(\Delta_t \mathbf{A})\mathbf{x} + \mathbf{A}(\Delta_t \mathbf{x}) = \Delta_t \mathbf{b} + \Delta_t \beta. \tag{49}$$

The ionosphere and the biases are assumed to not change significantly between time steps, so that

$$(\Delta_t \mathbf{A})\mathbf{x} + \mathbf{A}\cancelto{0}{\Delta_t \mathbf{x}} = \Delta_t \mathbf{b} + \cancelto{0}{\Delta_t \beta} \tag{50}$$

Using the definitions

$$\mathbf{D} \triangleq \Delta_t \mathbf{A} \tag{51}$$

$$\mathbf{c} \triangleq \Delta_t \mathbf{b} \tag{52}$$

the resulting equation can now be written in the same form as Equation 36:

$$\mathbf{D}\mathbf{x} = \mathbf{c}. \tag{53}$$

A centrally-weighted subset of measurements within a time window is used for the reconstruction. This increases the number of measurements, in order to counteract the underdeterminedness of the problem. Equation 53 can then be solved using data from time windows surrounding each time step, with each inversion giving estimated electron density distributions $\hat{\mathbf{x}}$.

Before inverting the problem, the electron density distribution $\mathbf{x}$ is mapped onto horizontal basis functions in the form of empirical orthogonal functions (EOFs) using the mapping matrix $\mathbf{X}$. This is done in order to further decrease the degrees of freedom and reduce underdeterminedness. The EOFs introduce a-priori information on the expected shapes of the electron density distribution and can be defined in both the horizontal and vertical dimensions. EOFs are basis functions generated from sets of normalised distributions, which can be based on ionospheric models or actual measurements. Spherical harmonic functions are sometimes used in the same way, especially in the horizontal dimension.

The mapping matrix contains the basis functions, and maps the scaling parameters $\mathbf{w}$ to the electron density by

$$\mathbf{Xw} = \mathbf{x}. \tag{54}$$

Because they can be mapped to the electron density, only the EOF scaling parameters $\mathbf{w}$ need to be determined directly. Using (54), the left side of (53) can be formulated as follows:

$$\mathbf{Dx} = \mathbf{D}(\mathbf{Xw}) = (\mathbf{DX})\mathbf{w} \tag{55}$$

so that the inverse problem can, with the notation

$$\mathbf{G} \triangleq \mathbf{DX}, \tag{56}$$

be written as

$$\mathbf{Gw} = \mathbf{c}. \tag{57}$$

Equation 57, in place of (53), is then inverted in two steps using Tikhonov regularisation (see section 3.1.1), with $\mathbf{R} = \nabla^2$, as

$$\hat{\mathbf{w}} = \mathbf{G}^{-g}\mathbf{c} \tag{58}$$

$$\hat{\mathbf{x}} = \mathbf{X}\hat{\mathbf{w}}. \tag{59}$$

Earlier versions of MIDAS used TSVD instead of Tikhonov regularisation to solve the inverse problem.

3.4 Summary

Computerised ionospheric tomography, or CIT, is a powerful tool used to image and study the ionosphere. CIT uses integrated measurements from instruments such as GNSS ground receiver networks and LEO satellites to reconstruct time-varying 2D or 3D images of ionospheric electron density. In order to solve the problem computationally the problem is discretised in voxels or layers of pixels, which in turn can be mapped to other basis functions to simplify the problem using a-priori information.

Tomography is at its core an inverse problem, and many techniques for solving inverse problems exist. In CIT it is often necessary to use some type of regularisation, such as Tikhonov regularisation or Truncated SVD. MIDAS, the algorithm used in this thesis, is implemented using Tikhonov regularisation and vertical EOFs as basis functions. MIDAS solves the inverse problem using sets of time-differenced input measurements in order to eliminate the need for bias estimation and integer ambiguity resolution.

4 | Detecting TIDs

Ever since the mid-20th century, a wide range of ionospheric instruments and techniques have been used to observe TIDs. Propagating wave features of TIDs were first detected by ionosondes in Australia by Munro (1948), who noticed wavelike perturbations appearing with time delays at separate ionospheric sounding locations. This led to studies using three-station triangulation to determine velocity and direction of the TIDs (Munro, 1950). Heisler, 1958, also in Australia, continued the work on TID observation using ionosondes and noted diurnal and seasonal patterns from a data-set spanning 1952-1955.

Another ground-based technique, the Incoherent Scatter Radar (ISR), was used by Thome (1964) to detect LSTIDs over Puerto Rico. Although ISRs are powerful instruments, they are complicated to build and expensive to run and therefore are only available in select locations. As more ISRs were constructed over the world this technique would, however, remain important to the study of TIDs, just like ionosondes.

Breitling et al. (1967) used 54 ionosondes worldwide to detect TID-induced perturbations in the F2-layer critical frequency attributed to high-altitude nuclear tests over Johnston Island in 1962. Hines (1967) also detected TID patterns in F2-layer critical frequency following the 1961 low-altitude Novaya Zemlya nuclear explosion.

Before GNSS, TEC could be measured using the Faraday rotation angle from satellites in orbit. Davis and Rosa (1969) used TEC observations from a geostationary satellite to detect LSTIDs and correlated them with geomagnetic activity. Using the same type of technology, Davis and Da Rosa (1970) observed TIDs in connection with the March 1970 solar eclipse. Lerfald et al. (1972) observed the same phenomenon with five ionospheric sounders but argued that there was not a strong case for a connection between the eclipse and the detected TIDs.

TEC from GNSS reference receiver networks became increasingly important to TID studies with the establishment of dense networks like the Japanese GPS Earth Obser-

vation Network (GEONET) used by e.g. A. Saito et al. (1998) and Yang et al. (2017). The emergence of GNSS has not only provided new tools to monitor TIDs but has also increased the interest in TID detection due to the demands for accurate ionospheric corrections in precise GNSS-based positioning.

A recently completed (July 2020) EU project, TechTIDE (`http://www.tech-tide.eu/`) focused on the real time detection of TIDs in Europe and southern Africa. The project uses eight different methodologies relying on networks of instruments like ionosondes and GNSS receivers to identify LSTIDs, MSTIDs and to indicate the level of ionospheric disturbances (Altadill et al., 2020).

This chapter introduces some important ionospheric observation instruments and their use in TID monitoring: ionosondes, ISRs, airglow imagers and GNSS, including the application of GNSS tomography for TID imaging.

4.1 Ionosondes

Ionosondes are radio wave instruments operating at HF that sound the ionosphere to profile the ionospheric electron density up to the local peak density. Ionosondes vertically transmit a series of signals at increasing frequencies and receive the reflections from the ionosphere. The travel time of a given signal indicates the height at which that frequency is being reflected. Reflection at vertical incidence occurs at the local plasma frequency, which is directly related to the electron density as given in Equation 17, repeated here for convenience:

$$\omega_p = \sqrt{\frac{N_e q_e^2}{\epsilon_o m_e}}, \qquad \text{(17 revisited)}$$

where $\omega_p = 2\pi f_p$ is the angular plasma frequency, ϵ_0 the permittivity of free space and q_e and m_e are the electron charge and mass. Thus, the ionospheric electron density N_e can be related to the signal travel time by the ionosonde, and therefore also to an altitude. No reflection happens at negative density gradients, so ionosondes can only detect electron density distributions that increase with height. This means that ionospheric layers have separate curves in ionograms (plots of the reflected frequencies against altitude) and that the decaying topside profile cannot be directly profiled and has to be inferred from models or other measurements.

An ionosonde station can measure TID signatures in the F-layer as fluctuations in the peak ionospheric electron density (e.g. Hajkowicz and Hunsucker, 1987). An example of a TID observation using the Roquetes ionosonde by Reinisch et al., 2018 is shown in Figure 6.

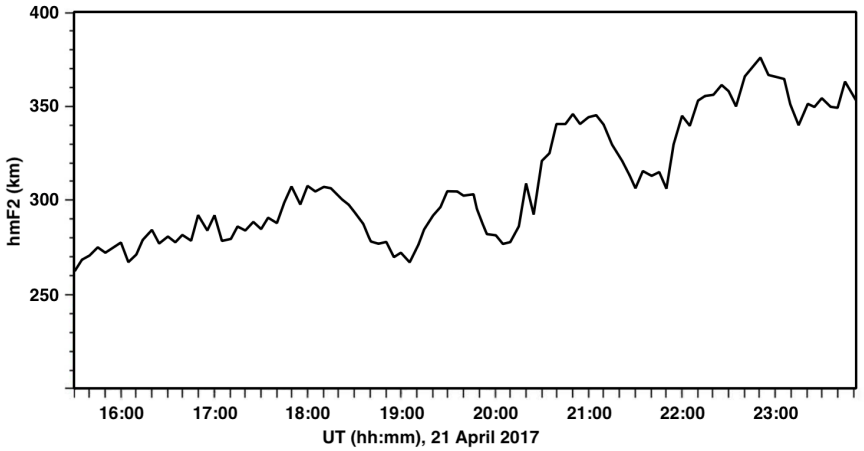

Figure 6: TID observed in hmF2 by an ionosonde in Roquetes, Spain (40.80°N, 0.50°E). Figure adapted from Reinisch et al. (2018).

Ionosonde-based TID-detection techniques are still improving and are often used in conjunction with other instruments. For example, Habarulema et al. (2018) combined ionosonde measurements with GNSS and satellite measurements to study LSTIDs. A new technique for TID detection that uses a network of ionosondes was introduced by X. Huang et al. (2016) which uses signals sent between different stations in a synchronised ionosonde network instead of vertical sounding. The method uses temporal variations in Doppler frequency and angle of arrival of received signals to estimate amplitude, wavelength, phase velocity and propagation direction of TIDs between the two locations.

4.2 Incoherent scatter radars

ISRs, first proposed by (Gordon, 1958), are powerful radars operating in the HF-UHF range able to use the incoherent backscatter from free electrons to determine electron density, electron and ion temperatures, and plasma velocity as functions of range (or altitude). The width of the measured incoherent backscatter spectrum is related to the thermal motion of the plasma as the scattered frequencies are Doppler shifted due to its motion. The power of the scattered signal is related to the scattering cross-section per unit volume and is therefore used to determine electron density. This backscatter is very weak, however, and requires a high-powered transmitter, a large antenna and a sensitive receiver. These requirements mean that ISRs are large-scale installations, and often set up as international collaborations in order to offset high construction and operation costs.

Examples of ISRs that have been used to study TIDs are The European Incoherent

Scatter Scientific Association (EISCAT) ISRs (e.g. Kirchengast et al., 1995; Vlasov et al., 2011; Van De Kamp et al., 2014; Nygrén et al., 2015) in Scandinavia and Svalbard, the Poker Flat ISR (PFISR) (Nicolls and Heinselman, 2007) in Alaska and the Arecibo ISR (Nicolls, Kelley et al., 2004) in Puerto Rico.

4.3 Airglow imagers

Another method employed to detect TIDs is the use of 630.0 nm band All-Sky Cameras (ASCs) that capture TID-induced perturbations in the *airglow* (e.g. Ogawa et al., 2002; Stefanello et al., 2015; Unewisse et al., 2015; F. Huang et al., 2018). Airglow is a faint emission of light in the upper atmosphere, which occurs as photons are released during recombination of ionised gas. It can therefore act as a tracer of the perturbations induced by TIDs. ASCs are however limited to nighttime operation and are sensitive to clouds blocking the field of view. In addition to the ground-based ASCs, there are examples of airglow imagers mounted on satellites used for observing TIDs, e.g. Rajesh et al. (2016). A series of airglow images showing several TID wavefronts from the study by F. Huang et al. (2018) are reproduced in Figure 7.

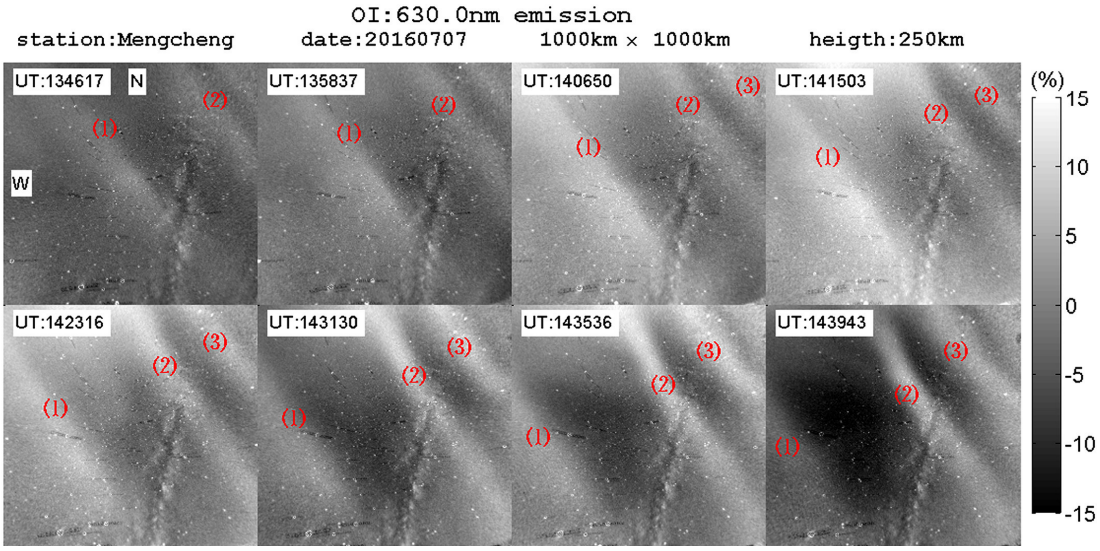

Figure 7: Series of 630 nm airglow images of a TID over Mengcheng, China (33.4°N, 116.5°E) on 7 July 2016. The colour denotes the deviation from a 1-h running average. Figure reproduced from F. Huang et al. (2018).

4.4 GNSS

Ionosondes and other ground-based instruments have nowhere near the geographic coverage of GNSS reference receiver networks, which has resulted in GNSS-derived TEC measurements being widely used for TID detection. Several different methods exploiting GNSS to map and detect TIDs have been used within the last decade, including GNSS tomography.

One of the first works to use GNSS to study TIDs over a large area was done by A. Saito et al. (1998), who used receivers in the dense GEONET reference network to construct 2D vTEC maps over Japan. These were de-trended by subtracting a 2 h running average to study nighttime TIDs. The study found that the method could discern TIDs, but the authors argued that complementary instruments like ISRs, airglow imagers and ionosondes would be needed to clarify the vertical ionospheric structure.

Hernández-Pajares, Juan and Sanz (2006) used a high-pass filter to de-trend TEC data and enable the detection of MSTIDs using cross-correlation. In order to estimate the spatial parameters, a system of equations was formulated for a network of receivers:

$$(\Delta \mathbf{r}_{ipp} + \mathbf{v}_{ipp}\Delta t) \cdot \mathbf{s} = \Delta t, \tag{60}$$

where $\mathbf{s}$ is the inverse of the velocity, or the slowness, of the TID wavefronts; $\mathbf{r}_{ipp}$ and $\mathbf{v}_{ipp}$ are the location and velocity of the IPP and Δt is the time travelled by the TID between two receivers determined by the maximum cross-correlation. Solving this system for s using, for example, the least squares method gives an estimate of the horizontal phase velocity and direction of the TID. An estimate of the period was also found by cross-correlation, using the time-lags for maximum (Δt_{max}) and minimum (Δt_{min}) correlation

$$T = 2|\Delta t_{min} - \Delta t_{max}|(1 - \mathbf{s} \cdot \mathbf{v}_{ipp}). \tag{61}$$

This cross-correlation approach was replaced by subtraction of the complex phase of the dominant Fourier transform terms in later work by Hernández-Pajares, Juan, Sanz and Aragón-Àngel (2012).

Spectral methods have been used by several other authors. For example, Katamzi et al. (2012), used the Fast Fourier Transform (FFT) algorithm on de-trended TEC to estimate TID periods from geostationary Faraday rotation observations. The de-trending chosen for this study was the subtraction a 6th order polynomial fit from the observed TEC. The authors also used a wavelet transform as an alternative method which, in addition to information on the wave period, indicates when specific periods are present in the

data.

Other works include Valladares et al. (2009), who used a de-trending technique similar to that in Katamzi et al. (2012), and imaged LSTIDs in TEC de-trended by subtraction of a 4th degree polynomial fit. The fact that TIDs produce large density gradients can be used to map large scale TIDs, and Borries et al. (2017) used plots of spatial and temporal TEC gradients taken from vTEC maps in order to identify LSTIDs during geomagnetic storm periods. In a recent thesis, Yang (2019) presented, tested and applied a multi-TID detection algorithm termed Atomic Decomposition Detector of Travelling Ionospheric Disturbance (ADDTID). In this approach, a predefined set of planar waves are used as a basis to find a best fit to de-trended vTEC maps. This allows the method to identify simultaneous - and even superposed - TIDs with different characteristics over dense GNSS receiver networks.

GNSS-based methods have also been developed to study TIDs in 3D. A method proposed by Kutiev et al. (2016) combines ionosonde and GNSS measurements by interpolating between full ionosonde profiles in order to construct a 3D electron density map for TID detection. The interpolated density profiles at each location are scaled by values from GNSS vTEC maps to provide a realistic spatial electron density distribution. This approach is similar to GNSS CIT in that it also produces 3D electron density maps from measurements.

4.4.1 TID imaging using GNSS CIT

Imaging TIDs using CIT is challenging due to some limitations inherent in common GNSS tomography methods. Most important is the lack of horizontal data paths. This deficiency is sometimes addressed by using some a-priori knowledge, such as using a basis set of EOFs seeded from ionosonde profiles or Chapman functions. The vertical structure of a TID, however, is not always well represented by Chapman functions, which may result in not having sufficient flexibility to accurately reproduce the TID-perturbed ionosphere.

This problem can be partly addressed by increasing the number of EOFs. Ssessanga et al. (2015) used three EOFs based on a modified version of IRI-2012 and a Multiplicative Algebraic Reconstruction Algorithm (MART) (Raymund et al., 1990) to image two TIDs over the dense GEONET receiver network in Japan. The results showed TIDs in 3D images with vTEC variations of ca 1 TEC. It was notable that certain specific wave shapes distinct to TIDs were visible in the reconstruction after de-trending by subtraction of a 1 h running average to isolate the perturbations in the images.

A different approach used by Chen et al. (2016) applied a varying regularisation constraint to force small gradients in the horizontal direction, but to allow large gradients near the F2-peak. Additionally, this method did not use an initial guess to iterate from, which is otherwise sometimes used to initialise an iterative solution (see e.g. Ssessanga et al., 2015; Raymund et al., 1990). The authors showed that the method was able to resolve a modelled MSTID with a 40 min period using simulated TEC measurements from GEONET.

4.5 Summary

TIDs can be detected using a variety of techniques. The most common are radio- and optics-based. The radio-based methods include ground-based techniques where radio signals are sent from ground and the measured response is used to characterise the ionosphere – ionosondes and ISRs – and the space based GNSS techniques. GNSS, when employed to study the ionosphere, exploits the dispersive effects imparted by the ionosphere on radio transmissions broadcast from GNSS satellites to receivers on the ground. Owing to dense networks of reference receivers, these have become one of the most useful instruments to image TIDs today. Some successful attempts have also been made to image TIDs using GNSS-Tomography over dense GNSS receiver networks. The optics-based instruments are called airglow imagers, and these are either mounted on spacecraft or on ground installations. Instead of using the interactions between the ionosphere and radio waves, these use the light emitted from the atmosphere in the form of airglow. TIDs perturb the atmosphere and leave traces in the airglow, which make them detectable by these instruments. This range of different instruments that are available provide good opportunities for verifying new methods of TID imaging and detection in areas where co-located data are available.

Part II

5 | MSTID effects on interpolation of ionospheric corrections

Commentary

This chapter examines the potential effects of TIDs on state-of-the-art ionospheric correction systems, and test certain improvements specifically designed for TID mitigation. Ionospheric corrections are commonly interpolated from values at reference stations, and this can introduce additional errors in the presence of TIDs and other sharp gradients.

The state-space representation GNSS correction framework used here was provided by Geo++ GmbH and the majority of the work was conducted during a secondment to Geo++ in Garbsen, Germany, in close collaboration with Francesco Darugna. A paper titled "*Mitigation of Severe Weather Events and TID Impact on the Interpolation of SSR Atmospheric Parameters*" by Darugna, Bolmgren et al. based partially on these results is currently under review, and this chapter reproduces figures from the same paper.

Simulations of TIDs approximated by simple harmonic waves were used to quantify the interpolation error as a function of wave amplitude and the receiver network density. This was done for a range of common interpolation methods, including several novel methods which take estimated TID parameters into account in order to reduce error. Following the simulations, the same methods were then applied to a TID scenario with a real data-set. The results showed that, in general, the application of TID-specific interpolation methods improve the interpolation when the network baselines are otherwise not sufficiently small.

An alternative solution is to forgo the interpolation entirely, and to instead calculate the corrections directly from a model of the ionosphere. Chapters 6-7 explore how tomographic imaging could be used to this end, and test how accurately such methods can reproduce TIDs.

5.1 Introduction

Various corrections are used in GNSS to get an accurate estimate of be observable pseudorange and carrier phase. While some navigation algorithms can be used independently of such corrections, rapid high-accuracy positioning commonly use corrections in order to obtain the fastest and most effective resolution of the phase ambiguity. As an example, using ionospheric corrections to improve ambiguity resolution is the main difference between traditional Precise Point Positioning (PPP) (Zumberge et al., 1997) (which uses the ionospheric free combination) and PPP–RTK (Wübbena, Schmitz et al., 2005). Ideally, TID-induced perturbations in the ionosphere should be included in the ionospheric corrections.

Pseudorange and carrier-phase corrections are sometimes supplied to the user either as observations from a nearby reference receiver, and such corrections are said to be in *observation space* representation. The other option is to provide separate corrections for each process affecting the observation, essentially one for each term in the observation equations. This paradigm is called *state space* representation (SSR).

This chapter concerns methods such as Network RTK (Wübbena, Bagge et al., 2001) (see section 2.4) that use reference network corrections in SSR. In N-RTK, ionospheric corrections are normally calculated for a network of reference stations and interpolated for the approximate user position. At the interpolation stage spatially small gradients in TEC, such as those induced by MSTIDs, can introduce additional errors on the order of the TID amplitude even if the TID should be accurately represented by the corrections estimated for each reference receiver. The potential interpolation errors for different interpolation methods and simulation scenarios are explored in section 5.3. In section 5.2 the interpolation methods are introduced, including methods adjusted to mitigate TID-induced interpolation errors. The results of each interpolation method for simulated TID scenarios are presented in section 5.3 and the results for a real data-set are presented in section 5.5. section 5.5 contains a discussion of the results and concludes the chapter.

5.2 Interpolation methods

The interpolation considered here is two-dimensional spatial interpolation. Three commonly used techniques were selected: Inverse Distance Weighted average (IDW), Ordinary Kriging (OK) and a Weighted Least Squares (WLS) fit of a planar surface. Wang et al. (2020) recently conducted a comprehensive study comparing several interpolation

techniques for PPP–RTK. The study found that, out of the tested techniques, a WLS fit of a low-order surface with distance dependent weights had the best adaptability. The WLS approach has therefore been selected for modification in order to better adapt it to TID mitigation, and these are tested addition to the established techniques.

5.2.1 Inverse distance weighted average

The IDW interpolation approach is a weighted average over the sample measurements, with the weights of the values decreasing with the distance from the query point. This weighting makes the method a natural approach to interpolation of GNSS corrections, since the distance dependant errors in GNSS have been shown to grow linearly with the length of the baseline (Beutler et al., 1988; Brunner, 1994; Schön, 2007). Mathematically, the interpolated value $f(\mathbf{x}^\star)$ at a query point $\mathbf{x}^\star$ is given by (Shepard, 1968)

$$f(\mathbf{x}^\star) = \sum_{i=1}^{n} w_i f(\mathbf{x}_i) \tag{62}$$

where $f(\mathbf{x}_i)$ are measured values, which in this case are the SSR ionospheric corrections at the reference receiver locations $\mathbf{x}_i$. The weights w_i are defined as

$$w_i = \frac{d_i^2}{\sum_{j=1}^{n} d_j^2} \tag{63}$$

where $d_i = ||\mathbf{x}_i - \mathbf{x}^\star||_2$ denotes the Euclidean distance between locations $\mathbf{x}_i$ and $\mathbf{x}^\star$. The summation in the denominator ensures that $\sum_{i=1}^{n} w_i = 1$, so that the estimation is unbiased.

5.2.2 Ordinary kriging

Kriging is a stochastic method of interpolation, which was originally developed for use on geological features in the field of geostatistics and is therefore well suited for spatial interpolation. OK is one of many kriging methods and is distinguished by the assumption of an unknown mean only in the neighbourhood of the query point. Like IDW, OK can be written as a weighted average on the form of Equation 62. In OK, however, the weights w_i are determined using *semivariograms* $\gamma(d)$ fulfilling

$$2\gamma(\mathbf{d}) = var(f(\mathbf{x} + \mathbf{d}) - f(\mathbf{x})) \tag{64}$$

where $\mathbf{d}$ is the distance vector between two points and $var(\cdot)$ denotes the stochastic variance. Semivariograms are used to express the different expected values for pairs of

points dependent only on the distance between them, rather than their locations. The functional form of the semivariograms is selected specifically for the dataset, and here an exponential semivariogram model was chosen:

$$\gamma(d) = \alpha d^{\beta} \tag{65}$$

where α is determined by least-squares fitting over each pair of data points. The exponent was set to $\beta = 1.5$ after some initial testing.

5.2.3 Weighted least squares planar fit

The WLS interpolation approach fits a planar surface to the measurement points, with weights proportional to the inverse distance from the interpolation location, to the negative power of p. The plane can be described mathematically by the equation

$$f(\mathbf{x}_i) = ax_{1,i} + bx_{2,i} + f(\mathbf{x}^{\star}) + \varepsilon_i \tag{66}$$

where the a, b and $f(\mathbf{x}^{\star})$ are the unknowns defining the plane and ε_i is the residual between the plane and the measurement at $\mathbf{x}_i = (x_{1,i}, x_{2,i})$. The three unknowns are estimated from a minimum of three measurements using weighted least squares with a diagonal weight matrix with elements d_i^{-p} for a selected power p, computed from the distances d_i between the measurement points and the query point. This ensures a de-weighting for values far away from the query point. The value of p determines the power of the weighting, with $p = 2$ being typical in GNSS applications. In this work, we also try $p = 4$ in order to examine the effects of a more aggressive weighting.

5.2.4 WLS with TID direction

The first modification of WLS to include TID information uses the direction of the TID, and we denote it WLSD. This rests on the idea that locations in the ionosphere that are affected by the same TID wavefront, i.e. are in phase, for example on the same wave-crest, are more similar than locations that are out of phase. With the assumption that the TID can be reasonably well described by a plane wave, the wavefront is a plane of equal phase perpendicular to the direction of propagation.

The WLS weighting is adjusted to give larger weights to measurements in areas perpendicular to the direction of propagation θ. This is implemented by replacing the distance d with a modified distance $\tilde{d}$

$$\tilde{d} = d\sqrt{1 + \cos^2\theta}. \tag{67}$$

The difference between using d and $\tilde{d}$ is visible when plotting isolines of the inverse distance weights. This is shown in Figure 8, where the yellow isolines corresponding to d form circles the while the blue isolines corresponding to $\tilde{d}$ form ellipses with semi-minor axis parallel with the TID direction of propagation. As in the implementation of regular WLS, this method was implemented for exponential powers $p = 2$ and $p = 4$.

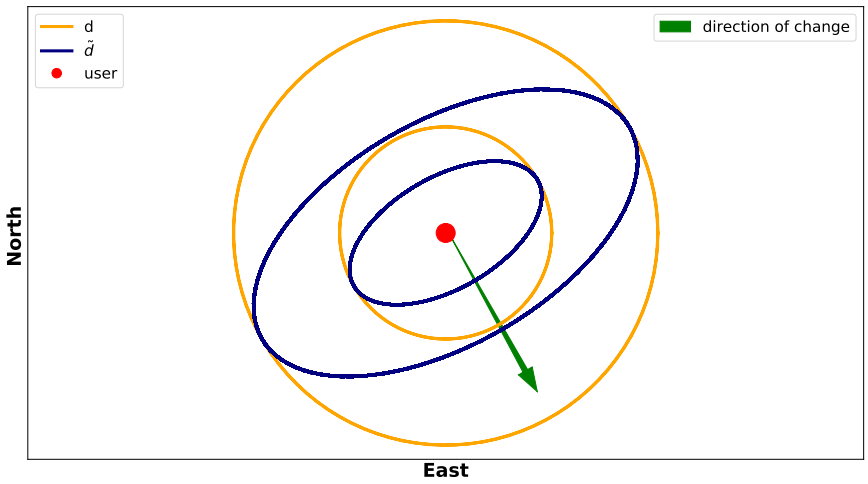

Figure 8: Weighting isolines. The green arrow indicates the direction of propagation of the TID. Reproduced from Darugna et al. (2020).

5.2.5 WLS with TID wavelength

The spatial scale of passing TIDs, in relation to the receiver density of the reference network, is important for determining the potential effect on the interpolation performance. In theory, the denser the network (or, the larger the wavelength,) the less need to apply additional weighting (as will be shown in section 5.3). To ensure stronger weighting for relatively small wavelengths, the power p can be made dependent on the ratio between wavelength λ and average network baseline b. We chose the function

$$\tilde{p} = p\sqrt{\frac{\lambda}{b}} \tag{68}$$

over $\tilde{p} = p\lambda/b$ in order to have a better behaviour for very small ratios. When this method is applied on its own, we denote it WLSL ($p = 1$). Both modifications, $\tilde{p}$ and $\tilde{d}$, can also be applied simultaneously, in which case we denote the technique WLSDL.

5.3 Interpolation of simulated data

5.3.1 Simulations

In order to evaluate the performance of the interpolation techniques with respect to TIDs, a simulated data set was generated. A simple plane harmonic wave was used to represent the passage of a TID and was injected into the ionospheric corrections before interpolation. The simulation is run through one cycle and is set to propagate in a westward direction. Real GPS satellite orbits were used, and both artificial and real reference station networks were used.

The artificial networks were constructed for two geometries: a circular and a square grid. The two artificial network geometries are shown in Figure 9, where the query point is shown as a red point in the centre of the networks. Three real network geometries were

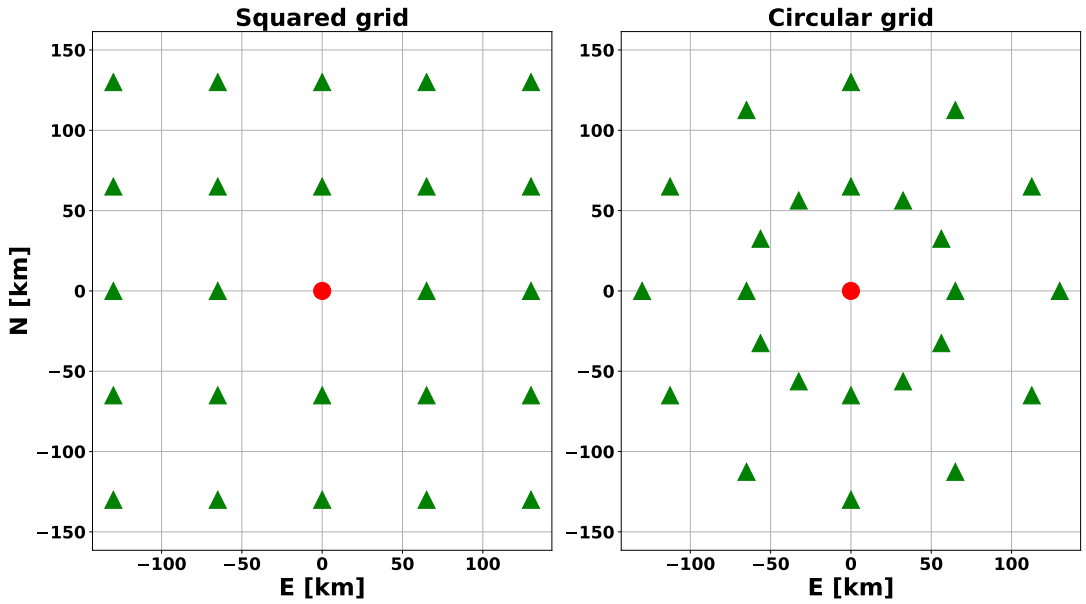

Figure 9: Two artificial reference receiver networks used for the interpolation simulations. The red point shows the location of the query point. Reproduced from Darugna et al. (2020).

added to give a more realistic picture of potential interpolation errors. These networks are subsets of the Netherlands Positioning Service (NETPOS) network, the Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) network in Germany and GNSS Earth Observation Network System (GEONET) network in Japan. The GEONET subset is centred on the prefecture of Okinawa, while the NETPOS and LGLN subsets cover the majority of the Netherlands and Lower Saxony respectively.

The respective subsets are plotted in Figure 10.

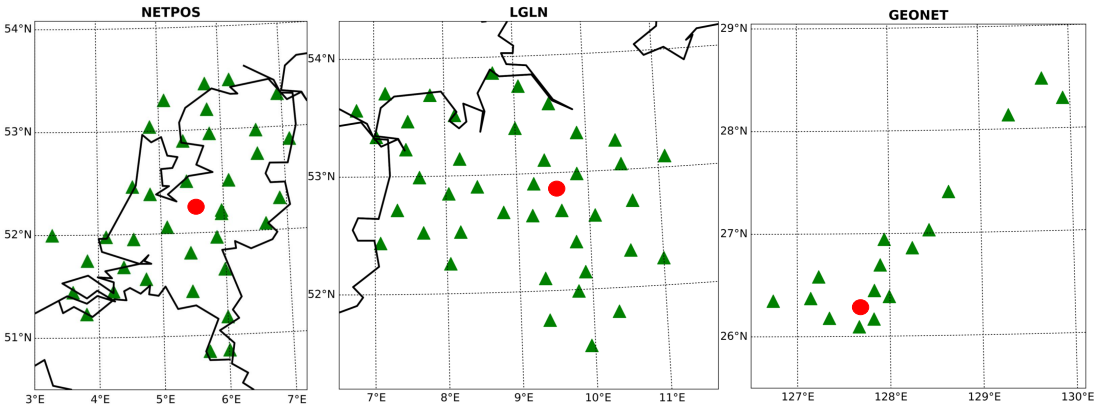

Figure 10: Real reference receiver networks used for the interpolation simulations. The red points show the locations of the receivers chosen as query points. Reproduced from Darugna et al. (2020).

5.3.2 Simulation results

For each network–TID combination, the simulation was run for one wave cycle and the maximum error at the query point throughout the cycle was selected. The ratio between this error and the modelled TID amplitude, A, is plotted on the y-axis in figures 11–15 against the wavelength/baseline ratio λ/b on the x-axis. Figures 11–12 show the artificial grid results, and figures 13–15 show the results from the three real network geometries. The numbers in the method abbreviations in these figures denote the value of p. For example, WLS2 implies $p = 2$.

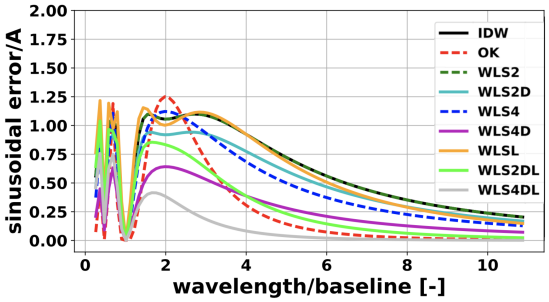

Figure 11: Square grid network simulation results. Reproduced from Darugna et al. (2020).

It is evident form each figure that smaller baselines, or larger wavelengths, in general decrease the interpolation errors for all methods. For $\lambda/b \lesssim 2$, however, the behaviour of some curves appears to break that trend. This behaviour, including maximum error

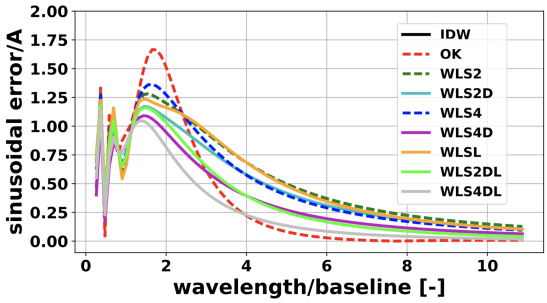

Figure 12: Circular network geometry simulation results. Reproduced from Darugna et al. (2020).

values at or close to zero, happen because of the regular geometry of the artificial networks. For example, the errors are zero in the square grid results (Figure 11) when the wavelength and baseline are equal, so that the values on all data points match the value at the query point exactly. When they are completely out of phase, however, the error increases beyond the TID amplitude.

By comparing the WLS2 (dashed green) to WLS2D (cyan), and WKL4 (dashed blue) to WLS4D (magenta) in the artificial network results, it is evident that the use of $\tilde{d}$ leads to smaller interpolation errors for both of the artificial networks. In these cases, the stronger power, $p = 4$, also shows smaller interpolation errors than $p = 2$. The use of $\tilde{p}$, however, does not show any improvement when not coupled with a directional approach. WLSL results in larger errors than WLS4 ($p = 4$), but when used with the directional methods (WLS2DL and WLS4DL), the error is below that of WLS2D and WLS4D. Out of the other reference methods, OK (dashed red) performs well for dense artificial networks ($\lambda/b \gtrsim 3$), while IDW (dashed black) performs on par with WLS2.

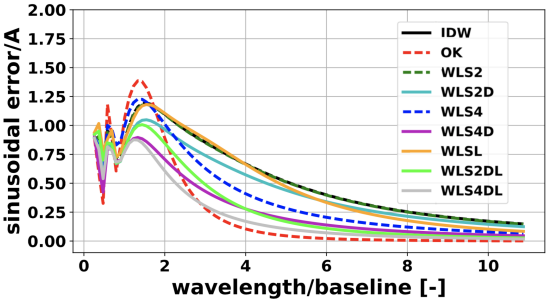

Figure 13: NETPOS subset network simulation results. Reproduced from Darugna et al. (2020).

The real network results generally follow the same pattern as the results for the artificial network. For $\lambda/b > 3$, OK performs best for both the NETPOS and LGLN networks,

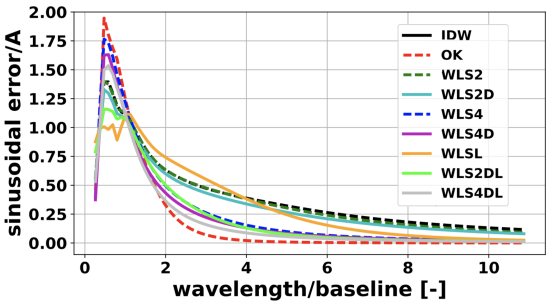

Figure 14: LGLN subset network simulation results. Reproduced from Darugna et al. (2020).

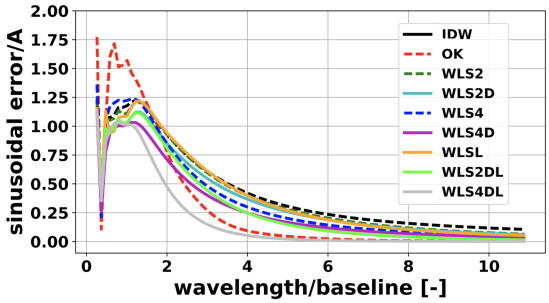

Figure 15: GEONET subset network simulation results. Reproduced from Darugna et al. (2020).

as in the circular network. The method including both $\tilde{d}$ and $\tilde{p}$ with $p = 4$, WLS4DL, on the other hand performs best for the GEONET subset, as in the results from the square grid.

5.4 Interpolation of real data

To verify the simulation results in real ionospheric conditions, an additional test with real SSR corrections was performed with the GEONET subset receiver network (see Figure 10). Ionospheric corrections were generated using GNSMART (`http://www.geopp.de/gnsmart/`) for a sample day, 14 May 2019, when TIDs were detected. Data from GPS, Galileo, GLONASS, and QZSS were processed in the generation of the corrections.

A particularly clear MSTID was detected in the TEC between 09:30-10:30, and this time period was selected for a case study. The TID direction and wavelength were estimated using a technique based on that developed by Hernández-Pajares et al. (2006), outlined in section 4.4 of Chapter 4. Instead of using a band-pass type filtering for the TEC series, the ionospheric model residuals were used as de-trended TEC, shown in Figure 16. These residuals are the measured TEC at the reference stations subtracted by the

spatial ionospheric model estimated in the GNSMART processing, at each time step. The ionospheric model has two parts – a global model and a satellite specific regional correction. The global model is a spherical harmonic expansion, while the regional model is a polynomial expansion. These residuals are also part of the ionospheric corrections and form the third part of the ionospheric corrections delivered to the rover.

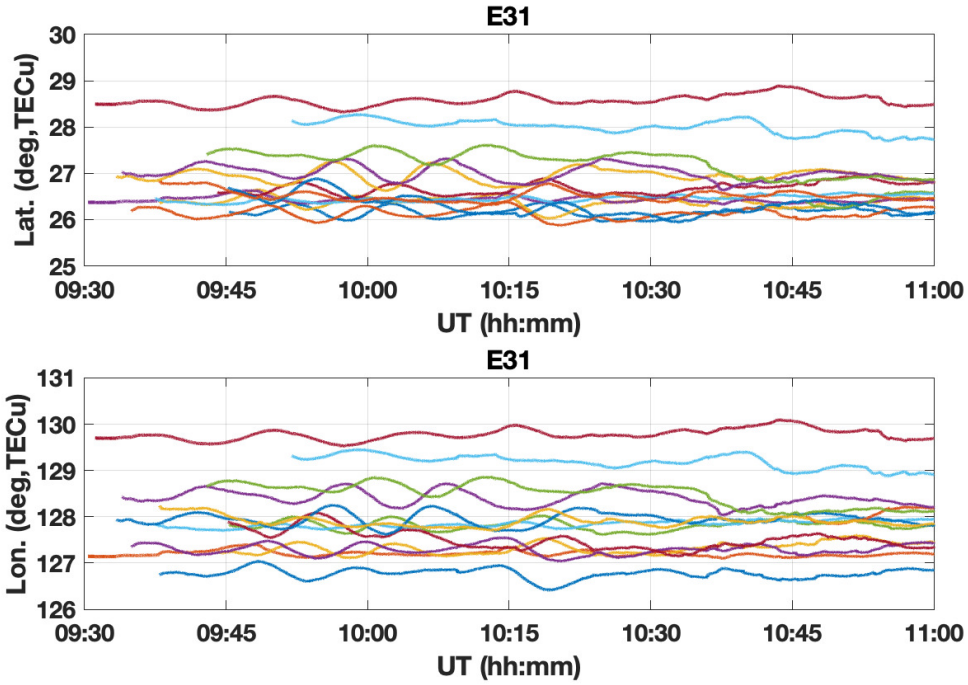

Figure 16: TEC values of the ionospheric grid residual for satellite E31 during the TID interval. Each line shows the perturbation around zero in TEC units, centred around the location of the receiver in degrees latitude (top) and degrees longitude (bottom). Reproduced from Darugna et al. (2020).

The estimated TID parameters are:

θ_{TID} $\approx 22°$ (N-E),
λ_{TID} ≈ 140 km,
T_{TID} ≈ 20 min,
A_{TID} ≈ 0.25 TECu.

The average baseline in the network is $b = 32\,\text{km}$, giving an estimated wavelength-baseline ratio of $\lambda/b \approx 4.4$. With these parameters, the results in section 5.3 suggest that the proposed modifications to the WLS technique should be able to improve the interpolation.

To test this, the performance of the WLS2DL and WLS2 techniques were used to interpolate the ionospheric corrections with one receiver excluded from the processing and used as query point. This was done for each station in the network, and the ratio between the average interpolation errors of the two methods for Galileo satellite E31 is plotted in Figure 17. In this figure, 9 out of 14 receivers show a ratio above one, indicating an improvement over the standard WLS2 method.

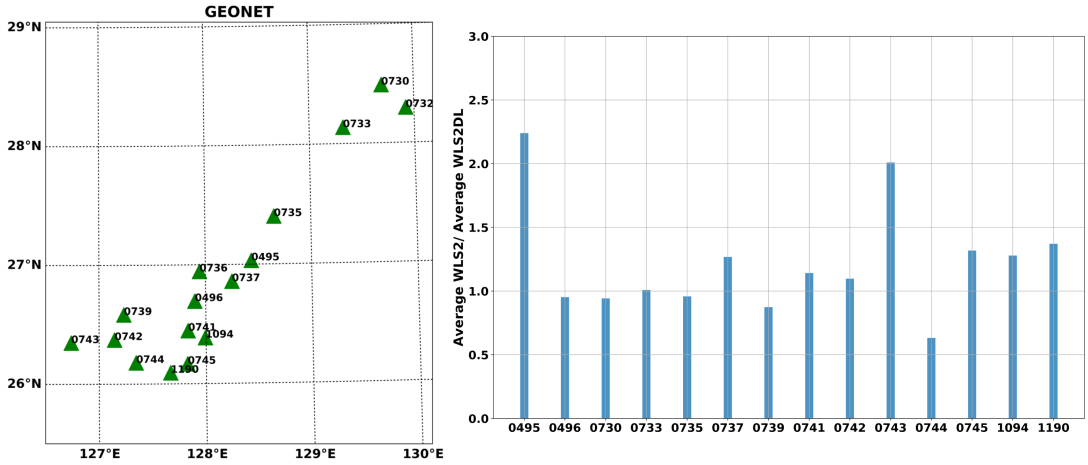

Figure 17: Ratios between average WLS2 and WLS2DL interpolation error (right) for each receiver location (left) of the Okinawa GEONET network. Reproduced from Darugna et al. (2020).

One test was also conducted for all methods considered in the simulation tests, for the same query point and reference used in the simulation test in section 5.3. The cumulative errors from this test are shown in Figure 18. Here, WLS2DL and OK performs the best, while WLS4DL performs the worst. WLS2, WLSL and IDW perform slightly better than WLS4, indicating that $p < 4$ is ideal in this case, contrary to the simulation results in section 5.3. WLS2D, WLS4D improve upon WLS2 and WLS4 respectively. For example, the number of absolute errors below 0.2 TECU changes from 75% with WLS2 to 90% with WLS2D. Using $\tilde{p}$ as well as $\tilde{d}$ improves this a further 6%, with WLS2DL.

5.5 Discussion and Conclusions

Large interpolation errors in ionospheric corrections can potentially hamper integer ambiguity resolution in N-RTK and PPP-RTK applications. Smaller interpolation errors, consequently, can decrease the time required for AR. TIDs pose a potentially

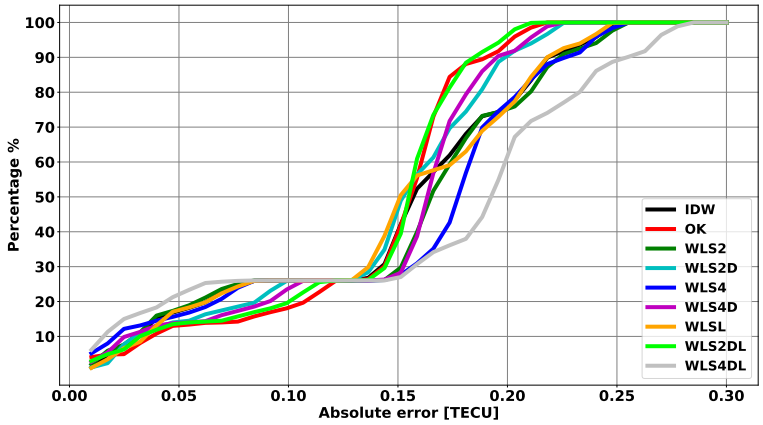

Figure 18: Cumulative error for satellite E31 interpolated for station 0741 (see left panel of Fig. 17). Reproduced from Darugna et al. (2020).

challenging interpolation problem, and this chapter presented simulations and a case study to compare the performance of different interpolation methods. These methods were inverse distance weighted average, ordinary kriging, and variants of a weighted least squares planar surface fit. Standard WLS was tested for exponential inverse weighting powers of $p = 2$ and $p = 4$. These two variants were in turn modified to take the direction of the TID, and the TID wavelength into account.

The simulation tests were conducted for five different geometries - two artificial networks and three real networks in Europe and Japan. The results showed that, as expected, a higher wavelength-network baseline ratio decreased the interpolation errors in all methods for $\lambda/b \gtrsim 2$. In the simulations, WLS with a weighting power $p = 4$ performed better than with $p = 2$. The use of directional information ($\tilde{d}$) led to smaller errors, as did the use of wavelength information ($\tilde{p}$) when combined with $\tilde{d}$. Wavelength-adjusted weighting power did not lead to significant improvements in the simulations when applied on its own.

A similar test was performed on a data set from an isolated subset of the Japanese GEONET network on 14 May 2019. A MSTID was detected, and its parameters were estimated using cross-correlation. Individual receivers were in turn used as rovers for the interpolation, and the use of both $\tilde{d}$ and $\tilde{p}$ improved the performance in 9 out of 14 locations.

These results show that using the estimated characteristics of TIDs in the interpolation of ionospheric corrections can help decrease interpolation errors. This has the potential to improve the reliability of ambiguity resolution for algorithms such as N-RTK and

PPP-RTK when TIDs are present. Estimated TID information (direction, wavelength, amplitude and average network baseline) could therefore be broadcast together with the SSR corrections in order to let the user adapt their interpolation strategy.

The directional interpolation method presented here has also been applied to counter sharp gradients in tropospheric signal delay during severe weather events. The direction of the gradient is then estimated using numerical weather models. More details on the tropospheric application can be found in Darugna et al. (2020).

In conclusion, this chapter has shown a potential use for estimated TID parameters in improving GNSS high accuracy positioning. Other estimation methods for TID parameters, or other base interpolation methods (e.g. IDW instead of WLS) could also be used. The generality of SSR corrections mean that these methods can be applied to different GNSS algorithms, not only N-RTK. It may also be possible to use external information from other atmospheric disturbances, not only TIDs, to achieve similar improvements in interpolation of GNSS corrections.

References

Beutler, G et al. (1988). 'Static positioning with the global positioning system (GPS): State of the art'. In: *GPS-Techniques Applied to Geodesy and Surveying*. Springer, pp. 361–380.

Brunner, FK (1994). 'On the deformation of GPS networks'. In: *Proc XX FIG Congress, Melbourne, March*. Vol. 501.

Darugna, Francesco et al. (2020). 'Mitigation of Severe Weather Events and TID Impact on the Interpolation of SSR Atmospheric Parameters'. In: *Advances in Space Research*.

Hernández-Pajares, Manuel, J Miguel Juan and Jaume Sanz (2006). 'Medium-scale traveling ionospheric disturbances affecting GPS measurements: Spatial and temporal analysis'. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–13. DOI: 10.1029/2005JA011474.

Schön, Steffen (2007). 'Affine distortion of small GPS networks with large height differences'. In: *GPS solutions* 11.2, pp. 107–117.

Shepard, D. (1968). 'A two-dimensional interpolation function for irregularly-spaced data'. In: *Proceedings of the 1968 23rd ACM national conference*, pp. 517–524.

Wang, Siyao et al. (Jan. 2020). 'A comprehensive assessment of interpolation methods for regional augmented PPP using reference networks with different scales and terrains'. In: *Measurement* 150, p. 107067. DOI: 10.1016/j.measurement.2019.107067.

Wübbena, G., A. Bagge and M. Schmitz (2001). 'Network-based techniques for RTK applications'. In: *GPS JIN*, pp. 14–16.

Wübbena, G., M. Schmitz and A. Bagge (2005). 'PPP-RTK: Precise Point Positioning Using State-Space Representation in RTK Networks.' In: *Proceedings of ION GNSS*. Pp. 13–16.

Zumberge, J. F. et al. (1997). 'Precise point positioning for the efficient and robust analysis of GPS data from large networks'. In: *Journal of geophysical research: solid earth* 102(B3), pp. 5005–5017.

6 | LSTID imaging by GPS tomography

Commentary

Our selected approach to detect and image TIDs was to use MIDAS – an established ionospheric tomography algorithm developed at the University of Bath. In order to examine if MIDAS GNSS tomography has potential for TID imaging, an initial study was conducted using the standard MIDAS configuration with GPS data. The scenario consisted of a TID with spatial features on the smaller end of the LSTID wavelength spectrum ($\lambda \approx 700$ km) but with large absolute amplitudes, surpassing one TEC unit. The LSTID occurred over North America on 31 October 2003 during the recovery phase of the 2003 Halloween Storm. The ionosphere was disturbed by the geomagnetic storm, making identification of TIDs among other perturbations potentially more challenging. The available GPS receiver network was also relatively sparse, which added to the difficulty of the inversion and made this an interesting initial test for MIDAS.

The analysis showed that, even in these challenging conditions, MIDAS was able to reconstruct the LSTID well enough for it to be identified using a sparse GPS TEC network. The results were compared to in situ electron density observations from the CHAMP satellite planar Langmuir probe, as well as peak height and cutoff frequency from the Dyess ionosonde. Both instruments suggested the presence of a TID. However, only three of the network receivers were located south of the Dyess station, and in this area the TID wave-pattern was not visible in the inversion, even if effects from the TID were visible in the CHAMP and ionosonde data in this region. This highlights the importance of GNSS ground-receiver coverage for reliable MIDAS tomographic inversions.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Tomographic Imaging of a Large Scale TID during the Halloween Storm of 2003			
Publication status (tick one) **Draft manuscript** ☐ **Submitted** ☐ **In review** ☐ **Accepted** X **Published** ☐			
Publication details (reference)	Bolmgren, K., Mitchell, C., Pinto Jayawardena, T., Bust, G., and Bruno, J. (2020). Tomographic Imaging of a Large Scale TID during the Halloween Storm of 2003, Ann. Geophys., in press, https://doi.org/10.5194/angeo-2020-26		
Copyright status (tick the appropriate statement) I hold the copyright for this material X Copyright is retained by the publisher, but I have been given permission to replicate the material here ☐			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: *75%* *Event selected by CM. Discussions with CM, TPJ and JB.* Design of methodology: *80%* *GB suggested using CHAMP PLP data.* Experimental work: *N/A* Presentation of data in journal format: *89%* *EM contributed OVATION Prime plot.*		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	

Tomographic Imaging of a Large Scale TID during the Halloween Storm of 2003

Abstract

The most intense ionospheric storm observed in recent times occurred between 29-31 October 2003. The disturbances to the high-latitude regions set off several Large-Scale Travelling Ionospheric Disturbances (LSTIDs), wavelike perturbations in the ionospheric electron density. This paper investigates one particular Travelling Ionospheric Disturbance (TID) on 31 October 2003 using North American Global Positioning System (GPS) receiver network data and a tomographic imaging technique. The TID has an estimated period of 30 min, an estimated horizontal wavelength of 700 km and propagates South-Westward over North America. The tomographic reconstruction of the wave is validated using a simulation of the observations and with independent observations from ionosondes and the CHAMP Planar Langmuir Probe. The results are discussed in the context of the magnetic and ionospheric conditions that may have contributed to the launch of the wave. Large-scale TIDs are challenging to study over large regions of the Earth, and the GPS network here is shown to offer a unique perspective on the spatial and temporal variation of the TID. The experimental results are backed up by simulations that show a denser network of receivers, as is available in more recent years, would produce improved accuracy in the TID imaging.

6.1 Introduction

Travelling Ionospheric Disturbances (TIDs) are ionospheric manifestations of Atmospheric Gravity Waves (AGWs) occurring in the neutral atmosphere (Hines, 1960). AGWs are buoyancy waves in the atmosphere and can be observed as TIDs when they transfer momentum to ions in the ionosphere by collision. Large Scale TIDs (LSTIDs) are a common occurrence during geomagnetic storms. LSTIDs are wavelike perturbations in the ionospheric electron density with typical wavelengths over 1000 km, periods between 0.5-3 h (Hocke and Schlegel, 1996) and typically travelling equatorwards from the auroral regions (Davis and Rosa, 1969). LSTIDs perturb the electron density and hence the Total Electron Content (TEC), the number of free electrons along a path through the ionosphere, on scales up to several TEC units (1 TECu = 10^{16} free electrons per m^2). TEC is proportional to the first order ionospheric delay of transionospheric radio waves propagating in the ionosphere and is therefore a crucial parameter for Global Navigation Satellite Systems (GNSS).

Between 29-31 October 2003 a series of Coronal Mass Ejections (CMEs) – expulsions of plasma from the solar corona – reached the magnetosphere of the Earth, causing strong geomagnetic storms. These are often referred to as the *Halloween Storm(s) of 2003*. The CMEs caused two sudden storm onsets on 29 October 2003 and 30 October 2003 (e.g. Mannucci et al., 2005; Horvath and Lovell, 2010). The planetary K-index (Kp) peaked at 9 on 29 and 30 October 2003, and 8 on 31 October. Kp remained above 4 throughout 31 October, which, although still disturbed, constituted the recovery phase corresponding to the second sudden onset. The Auroral Electrojet index (AE) reached a maximum of 1827 nT at 06:31 UTC on 31 October, which is plotted in Figure 19. Change in AE is related to auroral ionospheric current activity, which has been correlated with the appearance of TIDs at mid-latitudes (Hajkowicz, 1991; Hunsucker, 1982; Hocke and Schlegel, 1996; Lewis et al., 1996). These TIDs are thought to be launched by Joule heating of the atmosphere caused by increased ionospheric currents. High variability in AE occurred several times throughout 31 October, as seen in Figure 19. This variability in AE provides evidence for a potential TID generation mechanism being present.

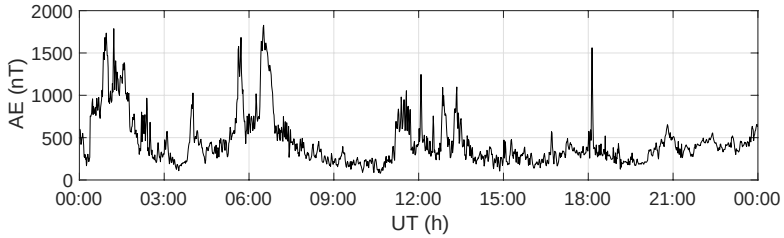

Figure 19: The AE index at 1 min intervals on 31 October 2003.

LSTIDs during the first two days of the October 2003 ionospheric storms have been studied extensively (e.g. Afraimovich et al., 2006; Ding et al., 2007; Perevalova et al., 2008; Valladares et al., 2009; Borries et al., 2009; Horvath and Lovell, 2010). This study focuses on the less intense third day of the storms, 31 October 2003, and specifically on a high-amplitude TID observed over North America in the local morning hours (16-20 UTC).

Section 6.2 covers the data instrumentation used for the study of the TID and shows examples of the GNSS slant TEC (sTEC) observed. In section 6.3.1, observations from different instruments and techniques - GPS tomography, an ionosonde and a space-borne Planar Langmuir Probe (PLP) are compared. To investigate the effects of using a sparse network of GPS receivers, an additional tomographic inversion using simulated data is performed in section 6.4. Section 6.5 contains a short discussion on the results and generation of the TID and final conclusions.

6.2 Data and instrumentation

The primary data used in this study were sTEC measurements derived from phase delay observations by a network of ground-based dual-frequency GPS receivers. In addition to the GPS sTEC used to image the TID, independent ionosonde data and measurements from the Challenging Minisatellite Payload (CHAMP) PLP were used to confirm the presence of a TID.

6.2.1 GPS TEC

The GPS receiver network is shown in Figure 20 and includes 40 stations in North America (listed in Table 2) which are part of the International GNSS Service (IGS) and UNAVCO networks.

Rx ID	lat (deg)	lon (deg)	Rx ID	lat (deg)	lon (deg)
alrt	82.4939	-62.34179	garl	40.4165	-119.3555
bake	64.3178	-96.00235	gtrg	43.2441	-113.2412
bcov	50.5443	-126.8426	guat	14.5904	-90.52018
bogt	4.65553	-74.10725	holm	70.7363	-117.7613
cags	45.5851	-75.80731	hvlk	37.6515	-99.10675
chur	58.7591	-94.08873	invk	68.3062	-133.527
cvms	35.5414	-89.64351	kely	66.9874	-50.94485
dsl1	70.3334	-148.4728	kuuj	55.2784	-77.74545
flin	54.7256	-101.978	mig1	34.0383	-120.3514
frdn	45.9335	-66.65992	mkea	19.8018	-155.456
modb	41.9023	-120.3028	sg00	47.9218	-97.08662
nain	56.537	-61.68864	thu2	76.537	-68.82508
ormd	29.2982	-81.10889	thu3	76.537	-68.82508
prds	50.8714	-114.2935	tono	38.0972	-117.184
ptal	49.2563	-124.8609	tukt	69.4383	-132.9944
qaq1	60.7153	-46.04779	vald	48.097	-77.56419
reso	74.6911	-94.8961	wslr	50.1265	-122.9211
ross	48.8337	-87.5196	yell	62.4809	-114.4806
sask	52.1963	-106.3984	ztl4	33.3797	-84.29673
sch2	54.8321	-66.83255	will	52.2369	-122.1679

Table 2: North American GPS receiver stations used for the tomographic inversion.

Slant TEC values were calculated using the geometry-free combination. It should be noted that MIDAS (section 6.2.1) uses time-differenced sTEC measurements, so satellite- and receiver biases which change slowly over time have no effect on the accuracy of the inversion (Mitchell and Spencer, 2003).

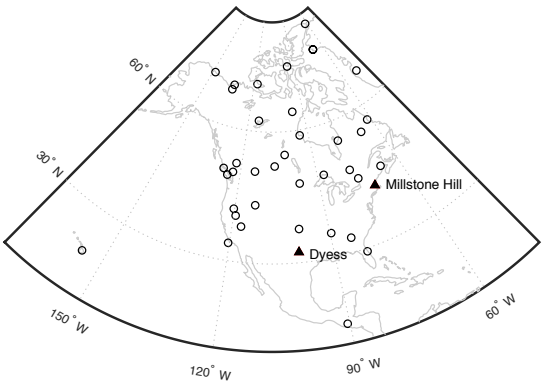

Figure 20: Network of GPS receivers used (circles) and the location of the Dyess and Millstone Hill ionosondes (triangles).

Figure 21 shows an example of pseudorange-calibrated sTEC observations from one receiver station, *tono*, where wavelike perturbations can be seen in the sTEC of several satellites. The satellites with the clearest TID signatures, PRNs 3 and 31, had Ionospheric Pierce Points (IPPs) moving north. It should be noted that the movement of the satellites relative to a TID may result in distortions to the apparent TID, as it introduces a Doppler-like shift in the apparent period of the TID perturbations (e.g. Wan et al., 1997; Hernández-Pajares et al., 2006; Penney and Jackson-Booth, 2015). Bolmgren et al. (2020) showed, using simulations, that MIDAS has the capacity to correctly image LSTIDs without explicitly taking this effect into account.

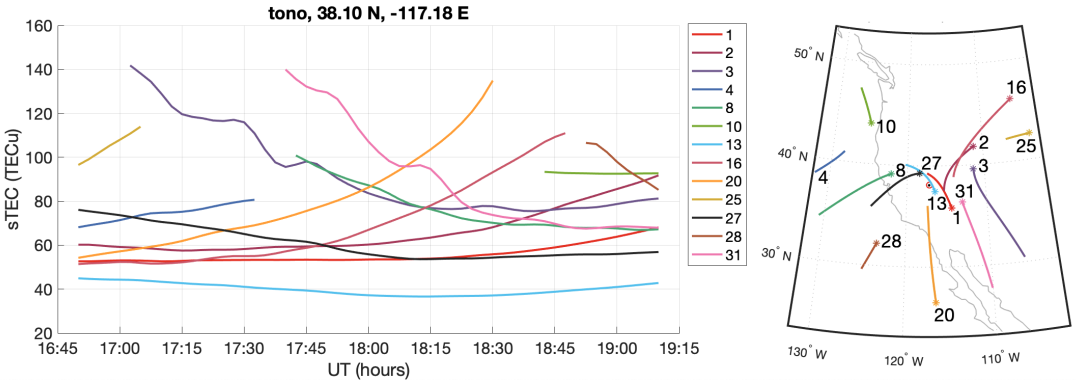

Figure 21: Biased sTEC from GPS receiver station tono on 31 October 2003.

MIDAS

Computerised ionospheric tomography is a method that can estimate the 2D or 3D ionospheric electron density over an area using integrated electron density measurements, such as TEC. In general, ionospheric tomography can be described as solving an inverse

problem formulated by the relationship between the geometry, the observations and the discretised electron density distribution. For a historical review of different methods of ionospheric tomography see Bust and Mitchell (2008).

In this study, the electron density was imaged using the Multi-Instrument Data Analysis Software (MIDAS) tomography algorithm (Mitchell and Spencer, 2003). MIDAS uses differential phase observations from a network of ground-based geodetic GNSS receivers and solves for an estimate of the ionospheric electron density. Empirical Orthogonal Functions (EOFs) are used as a change of basis in the height dimension; this constrains the problem by decreasing the degrees of freedom and by providing a basic structure to the variation of electron density with height. MIDAS has previously been tested as a TID imaging algorithm using a simulation approach in Bolmgren et al. (2020), which established that the algorithm can successfully reproduce LSTIDs using GNSS data. In this study we will show that this is possible with real data even in relatively challenging conditions.

6.2.2 Ionosondes

The first scientific observations of TIDs were made using ionosondes (Munro, 1948). Ionosondes are ground based radio instruments that characterise the bottomside electron density of the ionosphere. Ionosondes work by generating signal pulses that sweep through a span of frequencies. The pulses reflected back to the Earth from close to the zenith are used to estimate the height distribution of the plasma frequency, which is proportional to the square root of the electron density, directly above the ionosonde. The highest plasma frequency is usually found in the F2 layer and is denoted foF2. Since electromagnetic waves with frequencies above foF2 pass through the ionosphere, ionosondes provide no information on the electron density above the height of the F2 layer (referred to as hmF2).

Ionosondes at Dyess (32.4°N, 99.8°W) and Millstone Hill (42.6°N, 71.5°W) were both active on 31 October 2003. Figure 20 indicates the locations of these two ionosondes. The Millstone Hill ionosonde is used as a reference when setting up the MIDAS EOFs, while measurements from the Dyess ionosonde are used in section 6.3.2.

6.2.3 CHAMP planar Langmuir probe

The CHAMP satellite was active for ten years between 2000 and 2010 and was equipped with atmospheric and ionospheric observation instruments. CHAMP has a near circular polar orbit and had an altitude around 390 km at the time of the storm, which usually

would be in the topside of the ionospheric F layer. This study makes use of electron density data from the CHAMP PLP, a planar Langmuir probe which was used to measure in situ electron temperature as well as electron density in the front of the spacecraft every 15 s. Details on the CHAMP PLP can be found in McNamara et al. (2007).

6.3 Results

Sections 6.3.1, 6.3.2 and 6.3.3 present the results in terms of the tomographic GPS inversion, foF2 and hmF2 from the Dyess ionosonde, and CHAMP PLP in situ electron density respectively.

6.3.1 Tomographic inversion

Differential phase observations from the GPS receiver network were used with MIDAS to estimate the ionospheric electron density distribution on 31 October 2003. The reconstructions in MIDAS used voxels of $2° \times 2° \times 10$ km in latitude, longitude and height respectively, and time steps of 10 minutes. Two EOFs were generated using a set of Chapman profiles (Chapman, 1931), adjusted to fit the vertical profiles observed by the Millstone Hill ionosonde.

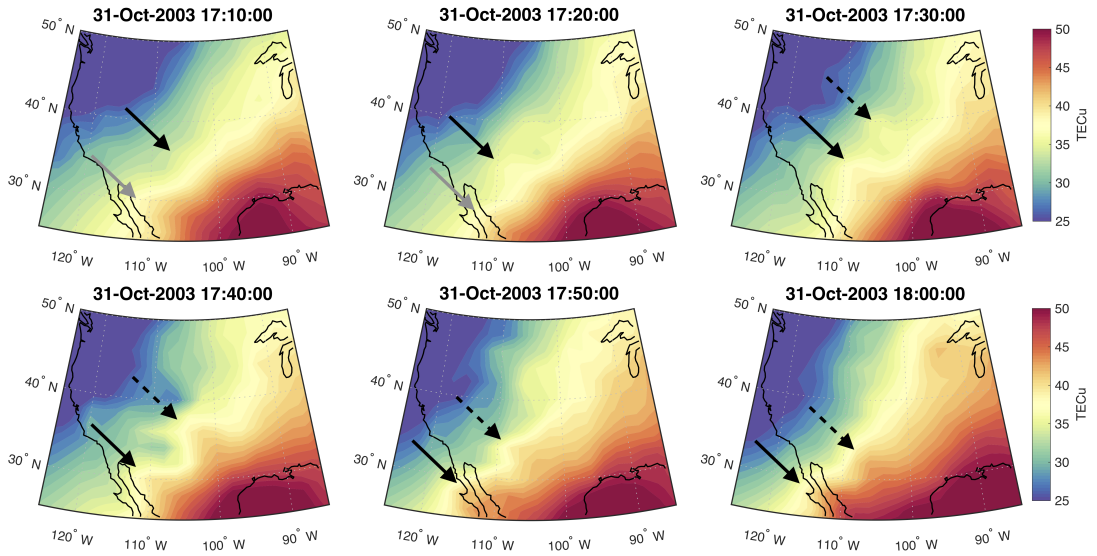

Figure 22: Series of vTEC from the MIDAS GPS inversion. Each frame is separated by 10 minutes. Arrows have been added to indicate identified wave-crests.

Figure 22 shows six consecutive time frames between 17:10-18:00 of the inversion results, with electron density integrated vertically to give vTEC. Between two and four wave fronts aligned NW-SE can be observed in the figure, spanning latitudes between 45°

and 30°. These features are also visible in the electron density viewed as a cross section spanning 100-1200 km in altitude along the direction of travel, shown in Figure 23. The wave-like perturbations are presumed to be the result of a passing TID.

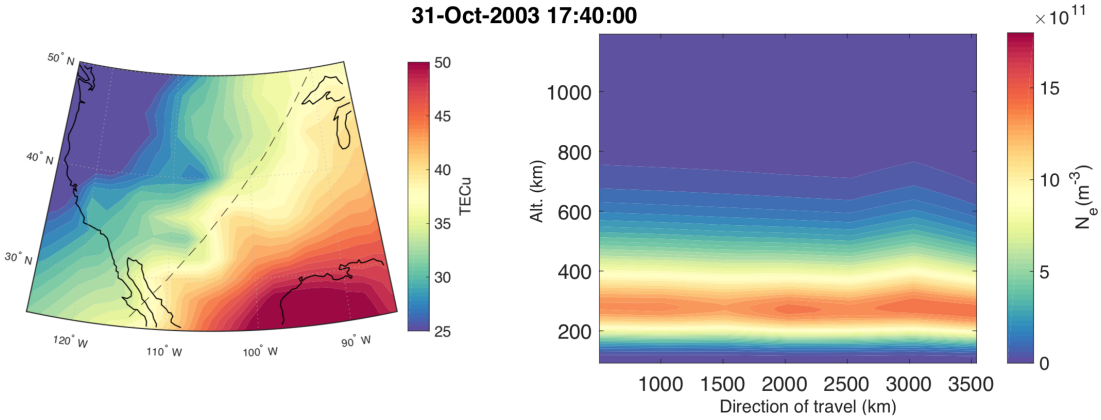

Figure 23: (Right) NE-SW cross-section of the full inverted electron density, and (left) the path of the placement of the cross-section in the vTEC map.

Using consecutive tomographic images from MIDAS, the TID parameters were estimated as follows: horizontal wavelength $\lambda_h \approx 700$ km, phase velocity $v_\varphi \approx 390$ m/s, and direction of travel $\approx 195°$ S-W . The period T was estimated as $T = \lambda_h / v_\varphi \approx 30$ min. These parameters would qualify the TID as medium scale, following the definitions in Hunsucker (1982). However, considering the high amplitude, geomagnetic conditions and equatorward direction of travel we will consider it a LSTID.

6.3.2 Ionosonde observations

The Dyess ionosonde is located within the area that was visibly affected by the TID in the MIDAS images. There is an indication of a periodical signature in the F2 layer critical frequency (foF2) with a 30 min period between 18:00 and 19:30 UTC, which may be related to the TID visible in the GPS data. However, the 15 min sampling makes it impossible to detect potential shorter period perturbations. In Figure 24, foF2 and hmF2 from the Dyess ionosonde are plotted against the equivalent parameters calculated from the MIDAS result. In Table 4 of Bruno et al. (2020), MIDAS results were compared against ionosonde data, and for a setup close to what is used here Bruno et al. (2020) found errors of 0.55 MHz in foF2 and 40 km in hmF2. The discrepancies in Figure 24 are on the same order. The other ionosonde with data readily available during this period, Millstone Hill (42.6°N, 71.5°W), does not show a similar indication of TID passage. This is expected, since it is located outside of the area visibly affected by the TID in the tomographic inversions.

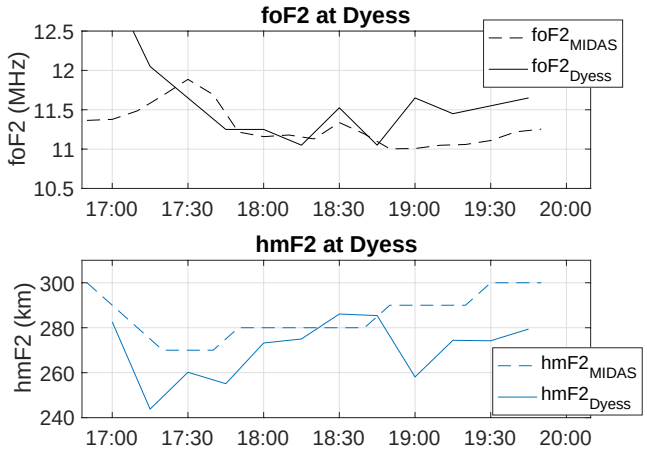

Figure 24: foF2 (top) and hmF2 (bottom) observations for the Dyess ionosonde (32.4°N, 99.8°W) and MIDAS equivalent sampled at the same location on 31 October 2003.

6.3.3 CHAMP PLP observations

The CHAMP satellite had one north-to-south pass over North America between 17:00 UTC and 19:00 on 31 October 2003, when the TID was visible in the GPS TEC. The in situ electron density measured by the PLP at altitudes between 391-395 km for this pass over North America is plotted in Figure 25. At 17:43 UTC, CHAMP passes North America at longitude 76°W, i.e. East of the area where the TID is visible in the tomographic images. Two dips in electron density separated by an apparent latitudinal wavelength of around 700-825 km are visible in Figure 25 between latitudes 15°N and 30°N.

The dotted line in Figure 25 shows the electron density estimated by MIDAS at 17:40 UTC, sampled at the location of CHAMP. Apart from not displaying the same wave perturbations, the electron density at this altitude is overestimated by approximately 3×10^{11} electrons per m^3. This is the result of a mismatch between the in situ observation and integrated estimate of the vertical density distribution in this area.

The perturbations in Figure 25 may indeed be caused by the passage of the TID seen further west in the tomographic images, but poor receiver coverage in the region may explain why the wavefronts do not appear to reach 76°W in the tomography result. The effect of possible poor data coverage is further examined by testing the tomography procedure on simulated data in section 6.4.

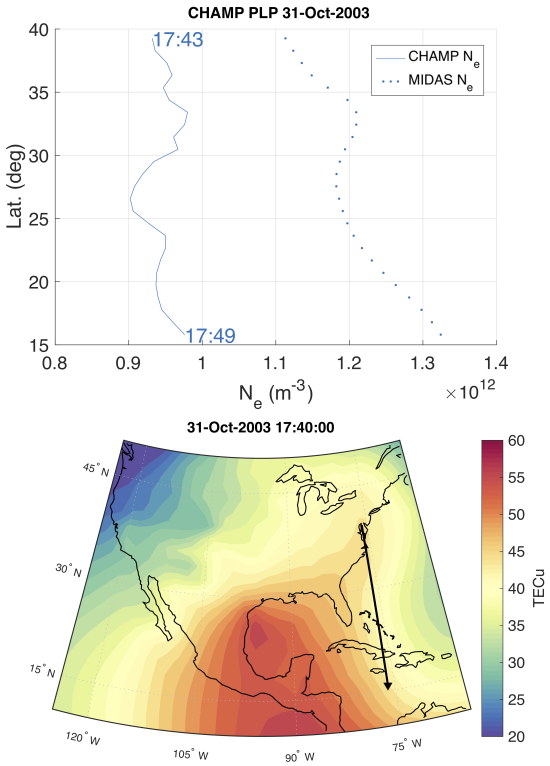

Figure 25: in situ electron density (top) measured by the CHAMP Planar Langmuir Probe (line) and the electron density sampled from the MIDAS inversion (dotted line) for 31 October 2003. Corresponding CHAMP satellite track is plotted on top of the MIDAS vTEC result for 17:40 UTC (bottom).

6.4 Method verification by simulation

The Dyess ionosonde and CHAMP PLP electron density both suggest the presence of a TID, but the wave-like features observed by these instruments are not clearly translated onto the same spatial and temporal coordinates in the MIDAS inversion results. The ionosonde suggests the presence of a TID with a period similar to that in the GPS inversion, but it appears later than it does in the inversion. The CHAMP satellite measurements suggest wave-like perturbations can affect the electron density as high up as 390 km in a region where the wave is not visible in the tomographic inversion. It is possible that these features are not visible in the tomographic images due to poor receiver coverage below 30 °latitude.

To investigate the effect of data-coverage and geometry used for the tomographic inversion, simulated TEC from a model ionosphere was inverted with MIDAS under the same geometric conditions (satellite geometry and receiver coverage) as the original

inversion. Any discrepancies between the model and simulated inversion results can be used to identify where there may be issues in the results presented in section 6.3.1. A second inversion of the simulated data, using a denser, fictional network is used to identify the effect of receiver geometry.

The TID parameters estimated in section 6.3.1 were used together with the Hooke (1968) TID model and the International Reference Ionosphere, IRI2016 (Bilitza et al., 2017), to generate a model ionosphere with TID, through which sTEC measurements were integrated (following Bolmgren et al. (2020)). A single frame of the model ionosphere is shown in Figure 26a.

The resulting inversion shows that the while the reconstruction with the regular network (Figure 26b) is able to conserve the main morphology of the TID, it does not correctly replicate the perturbations of the wave East of 105°W and South of 30°N. In addition, the wavefronts in Figure 26b appear skewed when compared to the model in Figure 26a. In all panels of Figure 26, a 1 h running mean was subtracted from each voxel post-inversion to minimise the background ionosphere and to better see the TEC perturbations caused by the modelled TID.

The wave is more accurately reproduced if a denser network of GPS receivers than was available in 2003 is used. Figure 26c shows the improved simulation result, which uses a larger number of receivers. The simulated receiver network is marked by points in the same sub-figure. This inversion more accurately reproduces the perturbations in Figure 26a, including the direction of the wavefronts.

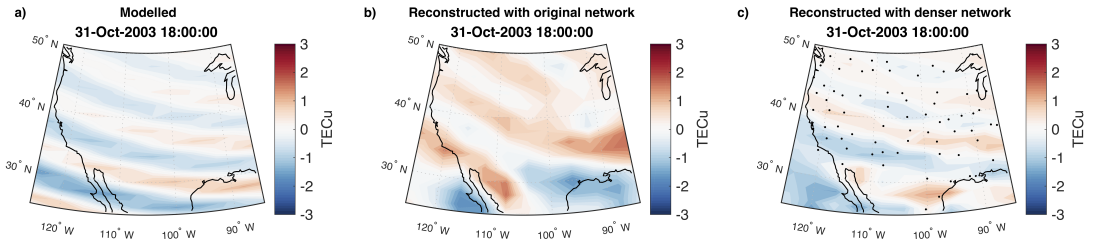

Figure 26: The modelled LSTID vTEC (a) and inversions from the simulated data (b) and (c) with a 1 h running mean background subtracted. Panel c uses a denser network than the real inversion, where the receiver locations of the network are indicated with black markers.

6.5 Discussion and conclusions

In this paper, we have used GPS tomography to reconstruct the ionospheric electron density over North America for 31 October 2003, the third day of the Halloween storm of 2003, and to identify a LSTID. The presence of a large-scale TID was evidenced by other instrumentation. A potential discrepancy in the TID morphology was observed between the measurements of two other instruments and the large-scale MIDAS reconstructions. While indications of the TID was captured by the Dyess ionosonde and CHAMP PLP, this was in areas where the MIDAS reconstruction showed no clear wave pattern. However, this was identified from computer end-to-end simulation to be the result of poor receiver coverage available for the MIDAS inversion, as discussed in section 6.5. The receiver network used has an approximate receiver density of 1 per $10^\circ \times 10^\circ$, compared to approximately 6 per $10^\circ \times 10^\circ$ for the denser synthetic network shown in Figure 26c. For comparison, the modern North American network used by Bruno et al. (2020) has an average receiver density close to 15 per $10^\circ \times 10^\circ$.

The observed TID had an estimated phase velocity of 390 m/s, an estimated period of 30 min, horizontal wavelength of 700 km and a southwesterly direction, suggesting a source in the auroral region. The high variability in AE occurring between 11:00 UTC and 14:00 UTC (Figure 19) may indicate a possible time of launch of the observed LSTID, if it were launched by Joule heating resulting from variations in the auroral electrojets. Another possible source mechanism may have been heating by auroral particle precipitation. The auroral oval was centred at latitude 63°N at this time with the region experiencing strong energetic particle precipitation at 14:30 UTC, as estimated by OVATION Prime 2013 (Newell et al., 2014) as shown in Figure 27. The highest levels of precipitation around the presumed launch time of the LSTID occurred between 08:00-10:00 Magnetic Local Time (MLT), which coincides with northern North America at the presumed launch time of the LSTID (11:00-14:00 UTC) and with the increased levels of AE activity around the same time. However, further analysis of additional datasets would be needed to obtain a detailed understanding of the generation mechanisms responsible for this LSTID. Since TIDs are effectively relative changes in the background electron density, the enhanced storm density likely contributed to the high perturbation amplitudes.

The work discussed in this paper built on that of Bolmgren et al. (2020), where MIDAS was demonstrated to be capable to image certain TIDs, and has shown that the tomographic algorithm is capable of imaging LSTIDs with relatively small spatial dimensions, provided that a sufficiently dense ground receiver network is available.

Tomographic maps like the ones produced here could be used in practical navigation

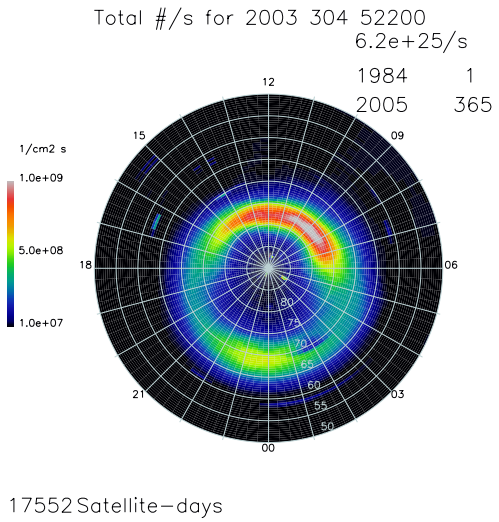

Figure 27: OVATION Prime 2013 particle flux for 31 October 2003 at 14:30 UTC.

systems to provide ionospheric delay corrections for GNSS positioning during LSTID activity, which may otherwise induce unmodelled TEC fluctuations impairing the quality of the solution.

Acknowledgments

This work was supported by funding from Horizon 2020 Marie Skłodowska-Curie Actions grant agreement No. 722023. CM acknowledges funding from NERC grant number NE/P006450/1. The authors thank the UML DIDBase (`http://umlcar.uml.edu/DIDBase`) for providing the data from the Dyess and Millstone Hill ionosonde stations. We thank the IGS (http://www.igs.org/) and UNAVCO (`https://www.unavco.org/`) for providing the GPS RINEX files, and the WDC for Geomagnetism, Kyoto (`http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html`) for providing the AE index.

References

Afraimovich, É L., É I. Astafieva and S. V. Voeykov (2006). 'Generation of ionospheric irregularities upon propagation of solitary internal gravity wave during the major magnetic storm of October 29-31, 2003'. In: *Radiophysics and Quantum Electronics* 49.2, pp. 79–92. DOI: 10.1007/s11141-006-0040-2.

Bilitza, D. et al. (2017). 'International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions'. In: *Space Weather* 15.2, pp. 418–429. DOI: 10.1002/2016SW001593.

Bolmgren, Karl et al. (Feb. 2020). 'Tomographic Imaging of Traveling Ionospheric Disturbances using GNSS and Geostationary Satellite Observations'. In: *Journal of Geophysical Research: Space Physics*. DOI: 10.1029/2019JA027551.

Borries, Claudia, Norbert Jakowski and Volker Wilken (2009). 'Storm induced large scale TIDs observed in GPS derived TEC'. In: *Annales Geophysicae* 27.4, pp. 1605–1612. DOI: 10.5194/angeo-27-1605-2009.

Bruno, Jon et al. (2020). 'A realistic simulation framework to evaluate ionospheric tomography'. In: *Advances in Space Research* 65.3, pp. 891–901. DOI: 10.1016/j.asr.2019.11.015.

Bust, Gary S. and Cathryn N. Mitchell (2008). 'History , Current State, and Future Directions of Ionospheric Imaging'. In: *Reviews of Geophysics* 46.2006, pp. 1–23. DOI: 10.1029/2006RG000212.

Chapman, S. (1931). 'The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth'. In: *Proceedings of the Physical Society*, pp. 26–45.

Davis, M. J. and A. V. da Rosa (1969). 'Traveling ionospheric disturbances originating in the auroral oval during polar substorms'. In: *Journal of Geophysical Research* 74.24, pp. 5721–5735. DOI: 10.1029/JA074i024p05721.

Ding, F. et al. (June 2007). 'Large-scale traveling ionospheric disturbances observed by GPS total electron content during the magnetic storm of 29-30 October 2003'. In: *Journal of Geophysical Research: Space Physics* 112.A6. DOI: 10.1029/2006JA012013.

Hajkowicz, L.A. (Aug. 1991). 'Auroral electrojet effect on the global occurrence pattern of large scale travelling ionospheric disturbances'. In: *Planetary and Space Science* 39.8, pp. 1189–1196. DOI: 10.1016/0032-0633(91)90170-F.

Hernández-Pajares, Manuel, J Miguel Juan and Jaume Sanz (2006). 'Medium-scale traveling ionospheric disturbances affecting GPS measurements: Spatial and temporal analysis'. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–13. DOI: 10.1029/2005JA011474.

Hines, C. O. (1960). 'Internal Amospheric Gravity Waves at Ionospheric Heights'. In: *Canadian Journal of Physics* 38.11, pp. 1441–1481. DOI: 10.1139/p60-150.

Hocke, K. and K. Schlegel (1996). 'A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995'. In: *Annales Geophysicae* 14.9, pp. 917–940. DOI: 10.1007/s00585-996-0917-6.

Hooke, W. H. (1968). 'Ionospheric irregularities produced by internal atmospheric gravity waves'. In: *Journal of Atmospheric and Terrestrial Physics* 30.5, pp. 795–823. DOI: 10.1016/S0021-9169(68)80033-9.

Horvath, Ildiko and Brian C. Lovell (Apr. 2010). 'Large-scale traveling ionospheric disturbances impacting equatorial ionization anomaly development in the local morning hours of the Halloween Superstorms on 29-30 October 2003'. In: *Journal of Geophysical Research: Space Physics* 115.A4. DOI: 10.1029/2009JA014922.

Hunsucker, Robert D. (1982). 'Atmospheric gravity waves generated in the high-latitude ionosphere: A review'. In: *Reviews of Geophysics* 20.2, p. 293. DOI: 10.1029/RG020i002p00293.

Lewis, R. V. et al. (1996). 'The generation and propagation of atmospheric gravity waves from activity in the auroral electrojet'. In: *Journal of Atmospheric and Terrestrial Physics* 58.6, pp. 807–820. DOI: 10.1016/0021-9169(95)00075-5.

Mannucci, A. J. et al. (2005). 'Dayside global ionospheric response to the major interplanetary events of October 29-30, 2003 "Halloween Storms"'. In: *Geophysical Research Letters* 32.12, pp. 1–4. DOI: 10.1029/2004GL021467.

McNamara, L. F. et al. (2007). 'Comparison of CHAMP and digisonde plasma frequencies at jicamarca, peru'. In: *Radio Science* 42.2, pp. 1–14. DOI: 10.1029/2006RS003491.

Mitchell, Cathryn N. and Paul S.J. Spencer (2003). 'A three-dimensional time-dependent algorithm for ionospheric imaging using GPS'. In: *Annals of Geophysics* 46.4, pp. 687–696. DOI: 10.4401/ag-4373.

Munro, G. H. (1948). 'Short-Period Changes in the F Region of the Ionosphere'. In: *Nature* 162, pp. 886–887. DOI: 10.1038/162680a0.

Newell, P. T. et al. (June 2014). 'OVATION Prime-2013: Extension of auroral precipitation model to higher disturbance levels'. In: *Space Weather* 12.6, pp. 368–379. DOI: 10.1002/2014SW001056.

Penney, R. W. and N. K. Jackson-Booth (2015). 'Mitigating satellite motion in GPS monitoring of traveling ionospheric disturbances'. In: *Radio Science* 50.11, pp. 1150–1164. DOI: 10.1002/2015RS005767.

Perevalova, N. P. et al. (Mar. 2008). 'Parameters of large-scale TEC disturbances during the strong magnetic storm on 29 October 2003'. In: *Journal of Geophysical Research: Space Physics* 113.A3. DOI: 10.1029/2008JA013137.

Valladares, C. E. et al. (2009). 'Simultaneous observation of traveling ionospheric disturbances in the Northern and Southern Hemispheres'. In: *Annales Geophysicae* 27.4, pp. 1501–1508. DOI: 10.5194/angeo-27-1501-2009.

Wan, Weixing et al. (1997). 'TID Observation Using a Short Baseline Network of Gps Receivers'. In: *Acta Geodaetica et Geophysica Hungarica* 32.3-4, pp. 321–327. DOI: 10.1007/BF03325503.

7 | TID imaging by tomography: method verification by simulation

Commentary

The standard mode of running MIDAS involves using GNSS relative TEC observations as an input in order to produce 3D images of the ionospheric electron density by tomography. Chapter 6 indicated that MIDAS can be used with GPS to image LSTIDs. To build on these results, it is interesting to see how well this method can be used to image TIDs of varying temporal and spatial scales. The analysis used synthetic TEC measurements generated from a modelled ionosphere. The modelled ionosphere was made up of a background, represented by IRI, and a foreground containing modelled TIDs, which were generated for sets of input parameters such as wavelength and wave period, ranging from MSTIDs to very large LSTIDs.

MIDAS is typically used with data from GNSS satellites in medium earth orbit. Due to the relative movement of these satellites and the TIDs, it may be beneficial to use satellites that appear stationary from Earth to avoid Doppler-shifted TID estimates, and satellites in GEO were therefore included in the analysis. It was shown that incorporating TEC data from geostationary satellites improved the results, specifically in the case of MSTIDs. LSTIDs were generally possible to image with data from regular GNSS orbits.

Using only geostationary measurements in the standard inversion approach resulted in images which showed only relative changes, while still reconstructing the TIDs. While this is an interesting and useful outcome, in practice, GEO satellites which can provide high-quality TEC estimations are not currently available in many parts of the globe (Kunitsyn et al., 2016).

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Tomographic Imaging of Traveling Ionospheric Disturbances Using GNSS and Geostationary Satellite Observations			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Bolmgren, K., Mitchell, C., Bruno, J., & Bust, G. S. (2020). Tomographic imaging of traveling ionospheric disturbances using GNSS and geostationary satellite observations. *Journal of Geophysical Research: Space Physics*, 125, e2019JA027551. https://doi.org/10.1029/2019JA027551		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: *75%* *Discussions with CM and JB.* Design of methodology: *80%* *Input from GB on TID model.* Experimental work: *N/A* Presentation of data in journal format: *100%*		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	

Tomographic Imaging of Travelling Ionospheric Disturbances Using GNSS and Geostationary Satellite Observations

Abstract

Travelling Ionospheric Disturbances (TIDs) are the manifestations of atmospheric gravity waves in the ionosphere. These disturbances have practical importance because they affect satellite navigation technologies such as Global Navigational Satellite System (GNSS), causing degradation in precise positioning applications. They also have scientific significance as their generation mechanisms and propagation are not fully understood. While there are specific instruments that can measure TIDs in certain locations there is a need for wide-area observations across extended geographical regions to continuously monitor their onset and spatial and temporal characteristics. This paper evaluates the use of observations from ground-based geodetic GNSS receivers to image TIDs using ionospheric tomography and data assimilation. Certain GNSS receivers also monitor signals from satellites in Geostationary Earth Orbit (GEO), which provide a unique perspective on the TID. The advantage of using the GEO data is investigated. A computerised simulation of GNSS observations is used for evaluation of the Multi-Instrument Data Analysis System (MIDAS) with GEO and regular GNSS geometry. The simulated observations are generated by integrating the electron density through a modelled TID-perturbed dynamic ionosphere between actual receiver and satellite positions. The output 3D electron density image series generated from the synthetic data by the MIDAS ionospheric tomography and data assimilation algorithm are compared with the input model ionosphere. Results show that GEO geometry improves the reconstruction of Medium Scale TIDs (MSTIDs) and smaller LSTIDs in cases where the movement of regular GNSS satellites in Medium Earth Orbit (MEO) may otherwise introduce distortions to the observations.

7.1 Introduction

The ionosphere, the ionised region of the upper atmosphere, has a significant effect on radio signals at frequencies up to several GHz. Radio transmissions are delayed and refracted when they encounter the electrically charged ionospheric plasma. This makes understanding the spatial and temporal distribution of the ionosphere important for radio communications.

Perturbations in the ionospheric electron density can have a significant impact on

precise positioning algorithms using Global Navigational Satellite Systems (GNSS) (Hernández-Pajares et al., 2006a; Hernández-Pajares et al., 2006b). Travelling Ionospheric Disturbances (TIDs) are a common type of ionospheric disturbance present at most latitudes. TIDs are the ionospheric signatures of Atmospheric Gravity Waves (AGWs), and manifest as propagating wave-like perturbations in ionospheric plasma density, velocity and temperature. The changes in plasma density are important for radio propagation. The amplitude, period, and spatial wavelength of TIDs vary, and they are typically categorised as either medium scale or large scale. Medium-Scale TIDs (MSTIDs) have periods below 1 h and horizontal wavelengths between 100 km and 1000 km, while large-scale TIDs (LSTIDs) have typical periods between 0.5 h and 3 h, and wavelengths above 1000 km (Hocke and Schlegel, 1996). TIDs can be initiated by processes in the troposphere as well as the ionosphere. An important source of MSTIDs is the solar terminator (Hernández-Pajares et al., 2006a; Nygrén et al., 2015). LSTIDs, on the other hand, are commonly associated with geomagnetically perturbed conditions and tend to travel equatorwards from auroral latitudes (Davis and Rosa, 1969). Although TIDs have been investigated for many years, interesting outstanding questions remain about the source mechanisms and it is still not possible to include them in ionospheric forecasting models.

TIDs have been observed by various ionospheric instruments. The first observations of TIDs where made by ionosondes in Australia, and were described by Munro (1950). Other commonly used instruments for TID observation include incoherent scatter radars, airglow imagers and dual frequency GNSS receivers. GNSS receiver networks, in particular, are important tools for studying TIDs over large spatial scales. Several studies have produced 2D maps of TID-induced perturbations in Vertical Total Electron Content (vTEC) over such networks, as was first done by Saito et al. (1998). However, GNSS observations of TIDs are subject to distortion due to the movement of the satellites relative to the TID, which must be taken into account in many cases (Wan et al., 1997; Van De Kamp et al., 2014; Penney and Jackson-Booth, 2015).

Each of these instruments has difficulty capturing the full spatial characteristics of TIDs without additional information. For example, Van De Kamp et al. (2014) used incoherent scatter radar observations to provide vertical information in order to complement the horizontal information provided by GPS observations. Incoherent scatter radars are, however, only available in a few locations worldwide, and in areas without complementing instruments like these it is still possible to infer horizontal information by using radio tomography methods.

Ionospheric radio tomography, introduced by Austen et al. (1988), is a method used to

image the ionospheric free electron number density $N(\mathbf{x}, t)$ distribution in 2D, 3D or 4D using sets of total electron content (TEC) data. TEC measured using GNSS is the electron density integrated along the ray path $\mathcal{R}_{ij}$ between satellite i and receiver j at time t,

$$TEC_{ij}(t) = \int_{\mathcal{R}_{ij}} N(\mathbf{x}, t)\, dl, \tag{69}$$

and is inferred from the phase advance and code delay observed by the receiver in the radio signals transmitted from the GNSS satellite.

Methods for ionospheric tomography were initially developed for low earth orbiting (LEO) satellites such as those part of the CICADA and TRANSIT constellations. These methods were successful in imaging MSTIDs along strings of ground receivers thanks to the fast LEO satellite passes (Pryse et al., 1995; Cook and Close, 1995; Kunitake et al., 1995). Increasing GNSS ground receiver coverage has since shifted focus towards GNSS satellites in medium earth orbit (MEO). A review of ionospheric tomography can be found in Bust and Mitchell (2008). GNSS tomography has also been applied in efforts to observe TIDs (as evidenced in Yizengaw et al., 2005; Lee et al., 2008; Ssessanga et al., 2015; Chen et al., 2016; Tang et al., 2016; Yin et al., 2017).

The movement of a GNSS satellite over a time period introduces a movement of the ray between satellite and receiver, and thus the integration path through the ionosphere $\mathcal{R}_{ij}$ in Equation 69. The movement of an ionospheric pierce point (IPP), a point where the ray intersects the ionosphere at a specified height, may reach velocities close to those of the TIDs themselves, which introduces distortions in the wave parameters observed in the TEC time series (Wan et al., 1997; Van De Kamp et al., 2014; Penney and Jackson-Booth, 2015).

A way to avoid these distortions is to eliminate the relative movement of satellite-receiver ray and TID by using observations from satellites in Geostationary Earth Orbit (GEO). Ground observations of Faraday rotation of signals sent from satellites in GEO have previously been used to study TIDs (e.g. Davis and Rosa, 1969). The advantage of using GEO has been recognised by recent work: for example, Huang et al. (2018) who compared observations of nighttime MSTIDs from GPS TEC, BeiDou GEO TEC and airglow observations over central China. They concluded that although the different observations generally agreed, GEO TEC and airglow were in closest agreement. They attributed this to distortion effects from GPS satellite movement.

In areas such as Europe and North America, which are outside the coverage of BeiDou satellites in GEO, satellites in the Satellite Based Augmentation System (SBAS) constel-

lations WAAS and EGNOS may be used instead. Kunitsyn et al. (2016) showed that the use of SBAS GNSS signals leads to a significantly higher TEC error when compared to signals from BeiDou satellites in GEO. However, Cooper et al. (2019) recently introduced a new method to calculate relative TEC from SBAS single frequency signals broadcast from GEO, improving the viability of using GEO TEC with ionospheric tomography in areas covered by SBAS. In summary, there are a number of GEO signals available that could be used in ionospheric tomography and have the potential to improve the resulting images in a range of geographical regions.

This simulation study examines the possible improvement in TID imaging from the addition of GEO TEC observations to the Multi-Instrument Data Analysis System (MIDAS) tomographic algorithm (Mitchell and Spencer, 2003). The simulation approach is outlined in Figure 28. The main steps are described in section 7.2. In section 7.2.1 an ionospheric electron density model is constructed as the ground truth. Section 7.2.2 explains how the model is used to generate synthetic TEC observations from satellites in GEO, and GNSS satellites in MEO. These are used by the MIDAS algorithm, described and referenced in section 7.2.3, to reconstruct the 4D electron density spatial and temporal distribution. In section 7.3 the MIDAS inversion results are compared to the modelled electron density in order to evaluate the inversion method and examine the effects of using observations from different satellite orbits to provide TEC.

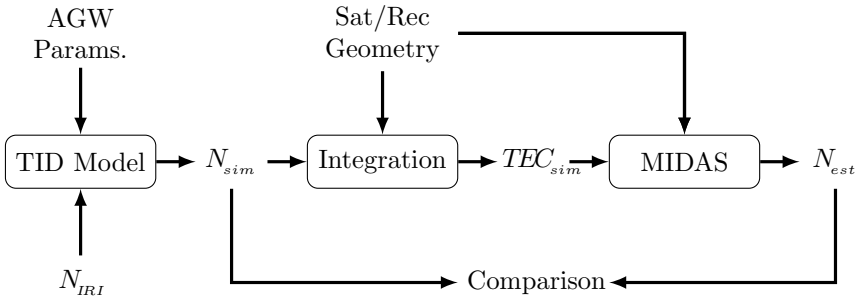

Figure 28: Flowchart of the method validation procedure. The Hooke TID model uses both AGW parameters and an IRI background N_{IRI} to produce the simulated test ionosphere N_{sim}. Receiver and satellite (GEO, GNSS-MEO or both) geometries are added first to integrate through N_{sim} to produce the synthetic TEC data TEC_{sim}, and then to the MIDAS algorithm for the reconstruction. Finally, the output N_{est} is compared to the input N_{sim}.

7.2 Method

7.2.1 TID model

A modelled ionosphere is used to evaluate the MIDAS algorithm for TID imaging by providing test data and a ground truth. The first part of the model is the unperturbed electron density, or background. This background is generated using the International Reference Ionosphere 2016 (IRI) (Bilitza et al., 2017). The TID-induced modulation of the background is calculated from AGW parameters using the physics-based Hooke (1968) model. In this model the electron density is given by

$$N_{sim}(\mathbf{x}, t) = N_0(\mathbf{x}, t) + \delta N(\mathbf{x}, t), \tag{70}$$

where the perturbation δN is given by

$$\begin{aligned} \delta N = N_0 \frac{U_b \sin I_b}{\omega} e^{k_{zi}(z-z_0)} \left[\left(\frac{\partial_z N_0}{N_0} + k_{zi} \right)^2 + \frac{k_{br}^2}{\sin^2 I_b} \right]^{1/2} \cdot \\ \cdot \cos \left[\omega t - \mathbf{k}_r \cdot \mathbf{x} - \tan^{-1} \left(\frac{k_{br} / \sin I_b}{\partial_z N_0 / N_0 + k_{zi}} \right) \right] \end{aligned} \tag{71}$$

where $\partial_z N_0$ is the vertical density gradient of the background, U_b is the neutral velocity in the direction parallel to the geomagnetic field at a reference height z_0, ω is the angular AGW frequency, I_b is the geomagnetic inclination, k_{br} is the real part of the wave vector $\mathbf{k}$ parallel to the geomagnetic field and k_{zi} is the imaginary part of $\mathbf{k}$, which is only nonzero in the vertical direction. The values for the geomagnetic field are taken from the International Geomagnetic Reference Field (IGRF) (Thébault et al., 2015).

This type of model has been implemented to model TIDs in several recent studies (Cervera and Harris, 2014; Nickisch et al., 2016; Mitchell, Rankov et al., 2017), and the model itself was tested by Morgan and Calderón (1978) who found relatively good agreement with measurements. The main drawbacks of this model are that there is no viscosity effect to damp out the wave at higher altitudes, and that the scale height is constant with altitude.

Here, following Hooke (1968), the assumption is made that there is no dissipation in order to get a simple estimate of the imaginary part of wave vector ($k_{zi} = 1/2H$). The horizontal and vertical wavelengths are related to each other and the angular frequency

ω by the AGW dispersion relation (Hines, 1960)

$$k_z = k_h \left(\frac{\omega_b^2}{\omega^2} - 1 \right) + \frac{\omega^2 - \omega_a^2}{c^2}, \tag{72}$$

where $\omega_b = (\gamma - 1)g/(\gamma H)$ is the Brunt-Väisälä frequency and $\omega_a = c/2H$ is the acoustic cutoff frequency. In these expressions, c is the speed of sound and γ the specific heat ratio. For simplicity, constant values H= 50 km, γ= 1.4 and $c = \sqrt{\gamma g H} = 829$ m/s are used.

The model Equation 72 was derived using a linearised perturbation treatment, and the physical approximations are valid only in the bottomside ionospheric F region. An additional exponential dampening is therefore applied above the F2 peak and below the upper F1 region when calculating the perturbation field $\delta N(\mathbf{x}, t)$. This approach may lead to an unrealistic response of the top-side ionosphere, where the effects of the TID are diminished.

7.2.2 Simulated TEC observations

The ionosphere simulation outlined in Figure 28 uses synthetic TEC observations TEC_{sim} generated from six different simulated ionospheric electron density structures $N_{sim}(\mathbf{x}, t)$. These model ionospheres are constructed as described in section 7.2.1, and only differ by AGW/TID parameter inputs to Equation 72, which are listed in Table 3. The background ionospheres were generated with IRI every 5 minutes, and interpolated linearly in time every 30 seconds, between 6 and 18 UT for 27 Sept 2016, a time with moderate geomagnetic conditions (Kp between 4 and 6) towards the end of the 24th solar cycle (monthly smoothed sunspot number of 33.2). The simulated TEC observations are integrated along each satellite-receiver ray path every 30 s according to Equation 69, with N equal to the modelled electron density, N_{sim}.

Table 3: TID wave parameters used to generate the six test ionospheres N_{sim}. $T = 2\pi/\omega$ is the wave period and $\lambda = 2\pi/k_h$ is the horizontal wavelength. All TIDs are directed North-South.

TID no.	1	2	3	4	5	6
T (min)	90	75	60	45	30	20
λ (km)	3000	2000	1500	1000	500	300

The 30 s interval is chosen as this is the usual rate at which GNSS data are collected and stored on ground-based geodetic receivers. It is assumed that the bending of the ray path due to refraction is negligible, which simplifies the ray paths to straight lines

between satellites and receivers. The coordinates of 191 receiver stations belonging to the International GNSS Service (IGS) and UNAVCO networks in North America, and real GPS and GLONASS precise satellite orbits provided by IGS are used to provide the geometry for the synthetic TEC observations. The positions of three satellites in the WAAS SBAS constellation: Inmarsat 4-F3 (98° W, PRN 133), Intelsat Galaxy 15 (133° W, PRN 135) and TeleSat Anik F1R (107.3° W, PRN 138) represent the geostationary satellite geometry. Any observation with elevation below 10° is discarded from the simulated synthetic observations. No Inter-Frequency Biases (IFBs) are added to the synthetic observations, as they have no effect on the regular MIDAS inversion (Dear and Mitchell, 2006; Bruno et al., 2020).

7.2.3 MIDAS tomography algorithm

The tomography method used in this study is the University of Bath MIDAS time-dependent inversion algorithm. In the ionospheric mode, MIDAS can take in any observation that is related to electron density. The standard mode for MIDAS to use GNSS is to take in dual frequency carrier-phase measurements to image the time varying ionospheric free electron distribution in three spatial dimensions. For these inversions Empirical Orthogonal Functions (EOFs) seeded with a set of Chapman functions (Chapman, 1931) provide the vertical electron density structure and reduce the complexity of the inverse problem. TEC observations, sampled from a time window weighted centrally around the current inversion time step, are used for the inversion. MIDAS can use either un-calibrated TEC that is ray-differenced for the inversion, as first done by Andreeva et al. (1992), or pre-calibrated TEC observations that have been corrected for biases. Ray-differenced TEC means the difference between consecutive 30 s TEC samples over the same satellite-receiver link. When pre-calibrated measurements are used, they must first be corrected for IFBs in both satellites and receivers. However, the calibration of IFBs can be problematic for real-time operations and hence MIDAS usually inputs only phase observations. The ray-differencing procedure is integrated into MIDAS, so the same data set can be used with either setting.

For this study 3D images of the electron density are reconstructed at 5-minute intervals. The data time window is 1 h wide and the voxels for the model and inversion are 25 km by 0.5° by 2° in height, latitude and longitude, respectively. The latitude-longitude grid is set to balance the capability to resolve TIDs while not creating too many unknowns to solve in the inverse problem. The data time window of 1 h was selected to provide enough data to be able to solve the inverse problem with this relatively high latitudinal resolution. This presents a trade-off between the size of the time-window and the spatial

resolution. A large time window lets more data be used for each inversion, but can cause other problems when the movement of the satellite and TID are involved. The height resolution is only used to map between the EOF representation of the height distribution and the voxel representation. This means that the height resolution can be chosen freely without affecting the number of unknowns in the inverse problem.

7.3 Results

Simulated line-integrated TEC observations were generated for the six different modelled TIDs in Table 3, generated as described in section 7.2.2, with two types of geometry: one using only GPS and GLONASS satellites in MEO, and a second using three satellites in GEO. It is important to stress here that the models are dynamic – they represent a propagating TID, and each integration to produce TEC is at the relevant time in the TID's movement. This is to make the simulated observations as realistic as possible. These TEC observations were then inverted using MIDAS to produce electron density structures $N_{est}(\mathbf{x}, t)$ for GEO satellite geometry, GNSS-MEO satellite geometry and a combined geometry and each of these for calibrated TEC and ray-differenced TEC input.

7.3.1 Modelling results

The modelling procedure described in section 7.2.1 was used to generate six TID-perturbed ionospheric electron density structures, N_{sim}, over North America. Figure 29 shows an example, with AGW/TID parameters from TID 4 in Table 3, of the two parts of the model: the background (figures 29a and 29b) and the perturbation (figures 29c and 29d). The full structure with background and TID combined is given in (figures 29e and 29f). The three top plots show slices through the three-dimensional electron density structures and the bottom plots show the same models integrated vertically to obtain vTEC.

7.3.2 Synthetic data results

TEC data were simulated for each model ionosphere, and all satellite geometries. Examples of simulated TEC observations from one receiver and six selected satellites (from a time-dependent ionosphere, a single frame from which is shown in Figure 29) are plotted in Figure 30 together with estimated wave periods T. Caution should be exercised here as the apparent wave periods when interpreted from this graph are dependent on the direction of movement of the satellite relative to the TID direction of

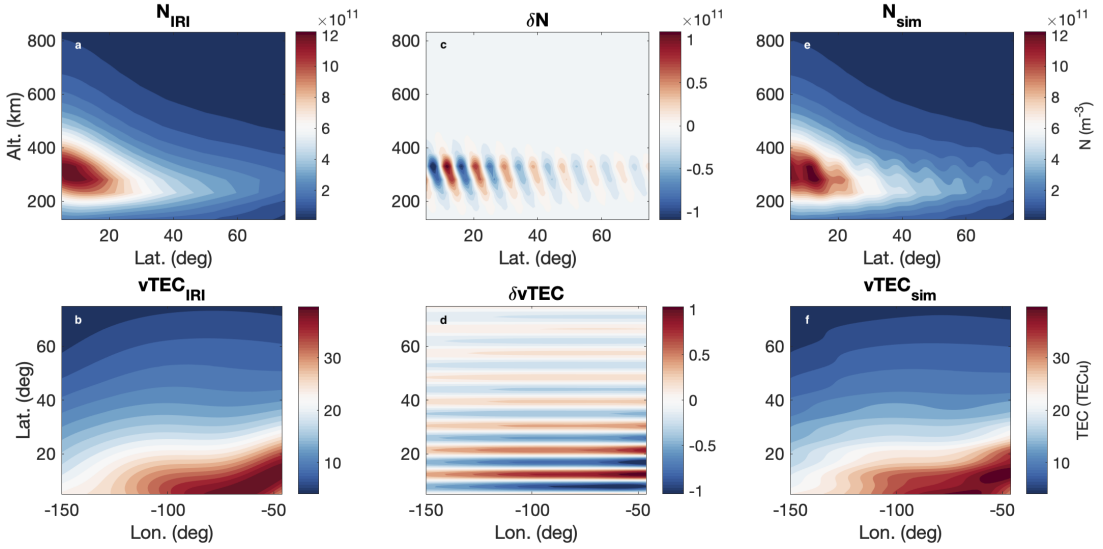

Figure 29: a and b show the background model ionosphere N_{IRI} as a cross section at 98^o W and as vertical TEC respectively. c-d show the modelled perturbation δN as generated by using Equation 72 with $N_0 = N_{IRI}$ and wave parameters corresponding to TID 4 in Table 3. When added, the background and perturbation form the simulated electron density N_{sim} shown in e-f.

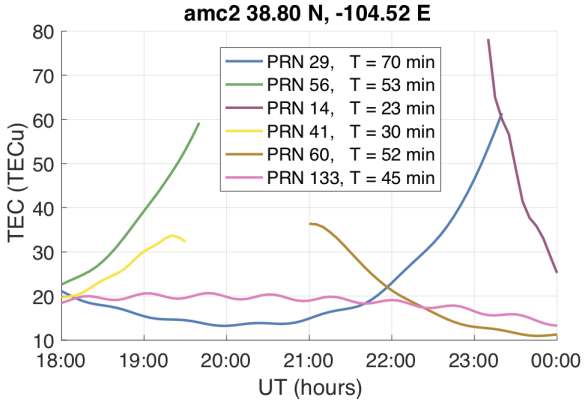

Figure 30: Simulated TEC measurements (TEC_{sim}) generated from a southward travelling TID (TID 4) with period 45 min and horizontal wavelength 1000 km for five selected GNSS-MEO satellites and one geostationary satellite (PRN 133) as seen from receiver amc2 (38.8° N, 104.5° W). The apparent periods T vary with the direction of satellite movement, shown in Figure 31

travel. This can be viewed alongside Figure 31 which shows IPP tracks at an altitude of 300 km. This altitude was chosen to coincide with the height where the modelled TID affects the largest effect on the ionosphere. For example, the satellite identified by pseudo-random-noise code (PRN) 41 moves against the TID propagation direction

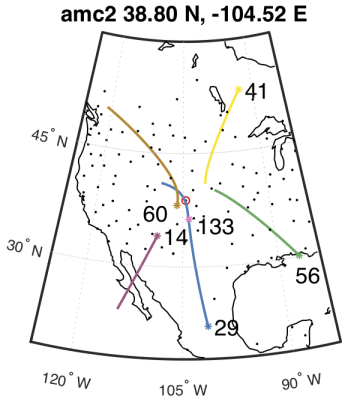

Figure 31: Tracks of satellite IPPs, at a height of 300 km, corresponding to the TEC time series in Figure 30. The last satellite positions as seen from the receiver are indicated by asterisks. The IPP of PRN 133 is shown only as an asterisk, as it appears stationary relative to the surface of the Earth. Black dots represent locations of receivers used in the inversion, and the location of amc2 is indicated by a circled dot.

and shows an apparent wave period shorter than in the simulation. PRN 60, on the other hand, moving more in the TID direction, displays a larger apparent period. The geostationary PRN 133 TEC_{sim} TEC series most closely reproduces the simulated TID period (45 minutes). It is clear from these examples that satellite movement can have a large impact on the apparent wave periods of the TEC. This shows that the observations themselves can exhibit distortion of the actual wave parameters.

7.3.3 Inversion results using pre-calibrated TEC

The first stage of evaluating the inversions is now to determine the difference between the simulated ionosphere and the MIDAS tomography electron density output when treating the synthetic TEC observations as 'perfect' (i.e. pre-calibrated) TEC input data. This means that no ray-differencing of the TEC is performed as part of the inversion. The MIDAS inversion results using GNSS-MEO and GEO geometry for the TIDs in Table 3 are shown in figures 32 and 33 as electron density cross section snapshots in height and latitude. The modelled wave perturbations are especially clear in the GEO inversion, and less evident in the GNSS results, where the waves are somewhat obscured by smaller artefacts. However, examination of the vTEC sampled at one point in the middle of the reconstruction grid, as plotted in Figure 34 reveals that the modelled TIDs are also present in the GNSS inversion.

In Figure 34 the modelled perturbations δN and the difference between the reconstruction and the modelled background ($N_{est} - N_0$) are compared in order to get a clearer view of

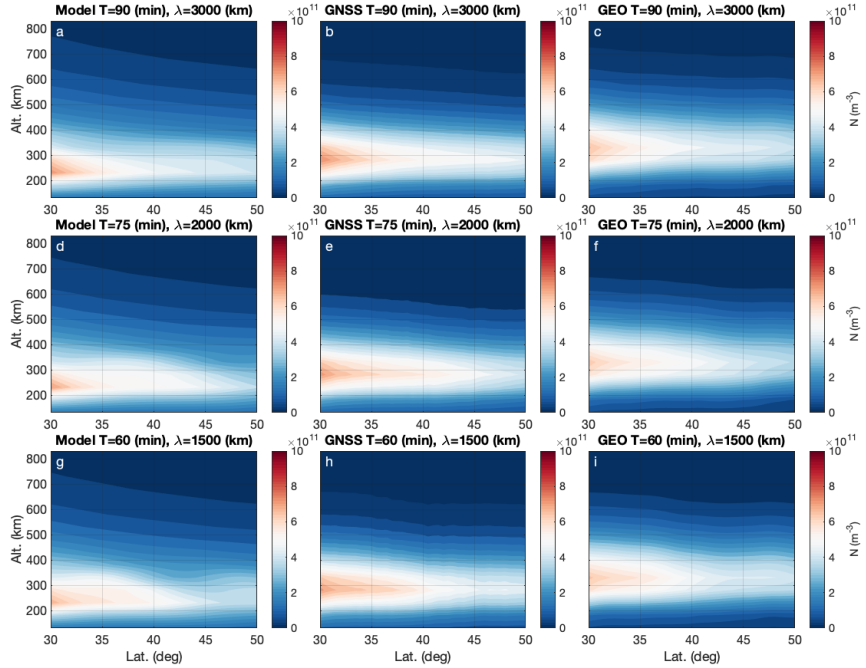

Figure 32: modelled electron density N_{sim} (a, d, g) and inversion results N_{est} (m^{-3}) (b, c, e, f, h, i) for TIDs 1-3 when treating the input data as pre-calibrated, i.e. without ray-differencing. b, e and h show GNSS-MEO results, and c, f and i show GEO-only results.

the wave-like behaviour. Figures 34a-f show the sampled vTEC as time series, and 34g-l show vTEC for all latitudes at one time step. Figures 34d-f and 34j-l illustrate how the GNSS inversion has trouble with smaller TIDs, while the GEO inversion vTEC is reasonably accurate for all but the smallest modelled TID (34f and 34l). It should also be noted that the background N_0 would not be known a-priori in a real data scenario and must be estimated by image filtering.

Some spatial artefacts smaller than the modelled TIDs appear in all GNSS inversions, as evident in Figure 34. For TIDs 1-2 they appear as small variations below 0.1 TECu in amplitude around an otherwise well estimated vTEC. They are more apparent for TIDs 3-6 in figures 34i-l, where the artefacts reach amplitudes around 0.2 TECu. It is interesting that the GNSS reconstructions tend to overshoot the TID wave amplitudes for the smallest TIDs by as much as approximately 5 times the amplitude, as is the case for one wave crest in Figure 34e.

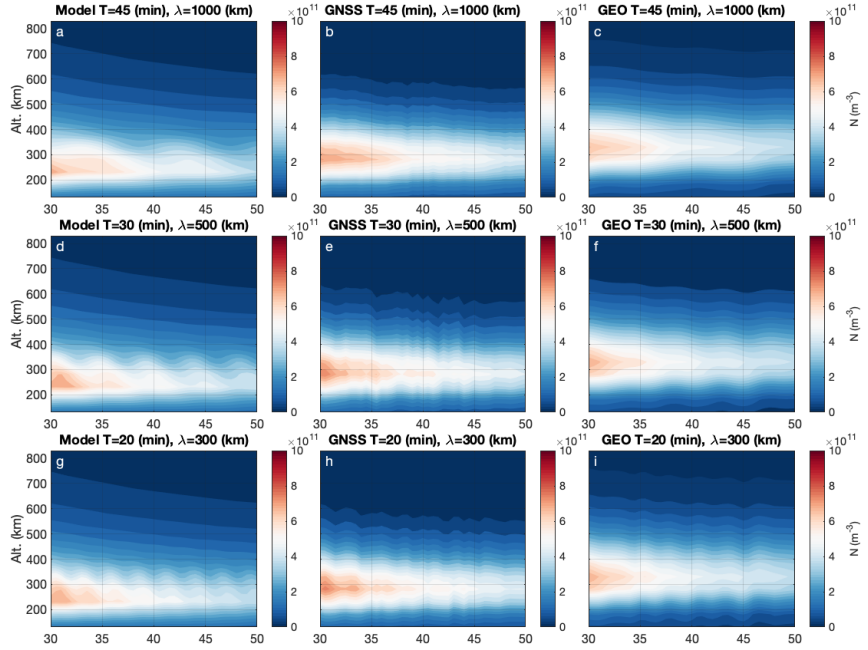

Figure 33: modelled electron density N_{sim} (a, d, g) and inversion results N_{est} (m^{-3}) (b, c, e, f, h, i) for TIDs 4-6 when treating the input data as pre-calibrated, i.e. without ray-differencing. b, e and h show GNSS-MEO results, and c, f and i show GEO-only results.

7.3.4 Inversion results using ray-differenced TEC

If no reliable IFB estimates are available, the MIDAS algorithm can run in a ray-difference mode, where the inversion algorithm inverts for changes in electron density rather than absolute values.

For slowly changing electron density N_{sim} the simulated ray-differenced TEC values become very small when using GEO satellite geometry, which makes this inversion configuration less sensitive to slow, large scale changes like those present in IRI. This means that GEO with MIDAS in ray-difference mode is much more sensitive to the sharp density gradients in the perturbation δN caused by the modelled TIDs than to the background N_0. This sometimes gives us a clear view of the modelled TIDs as displayed in Figure 36, which shows the TEC inversion results for TIDs 4-6. For the larger TIDs 1-3 in Figure 35, the structures are faint compared to the modelled density perturbations. This effect is also seen in the vTEC of Figure 37, where larger TIDs yield underestimated amplitudes when using GEO geometry.

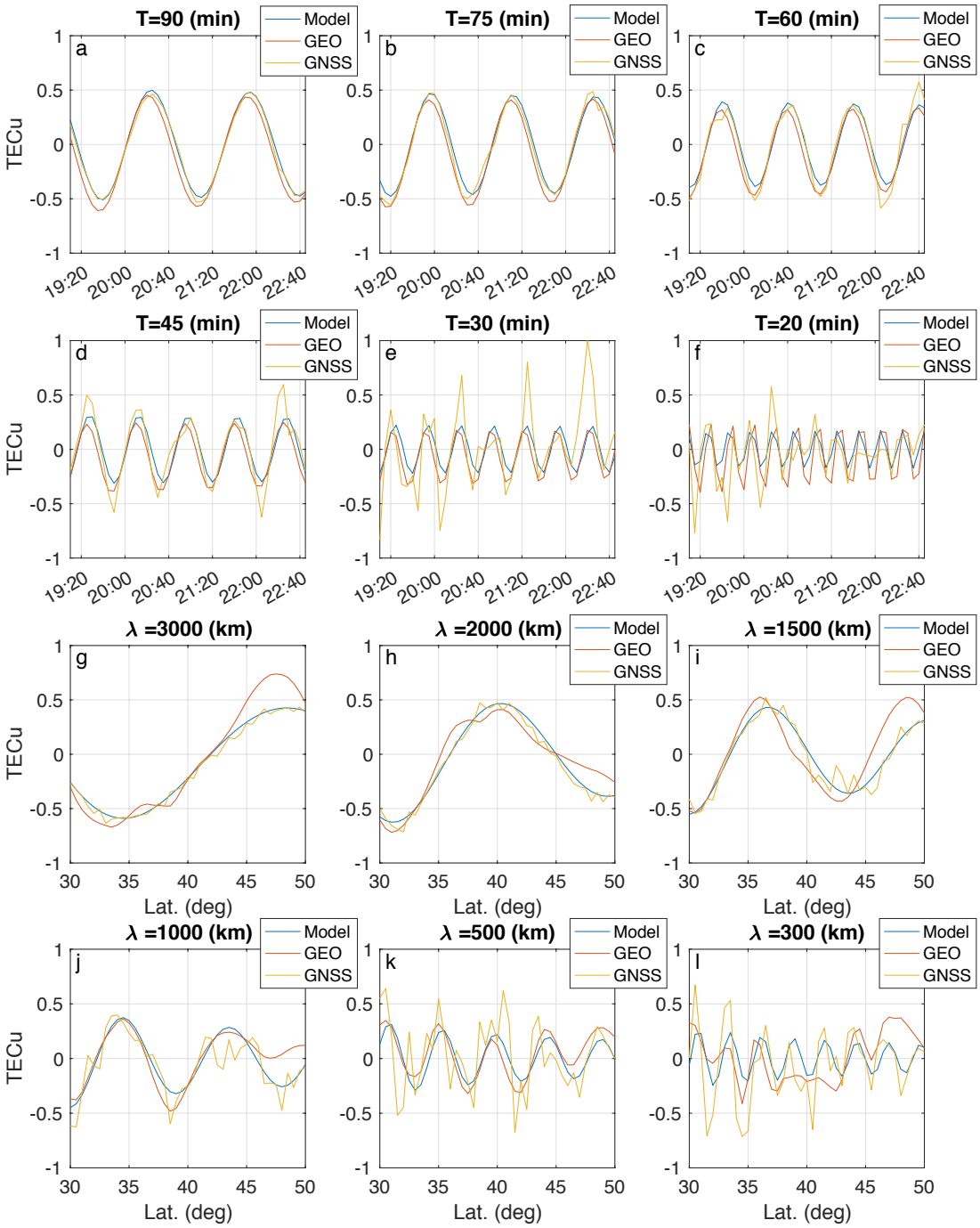

Figure 34: Pre-calibrated TEC inversion results ($vTEC_{est}^{GNSS-MEO} - vTEC_0$) and ($vTEC_{est}^{GEO} - vTEC_0$) plotted against the model δTEC.

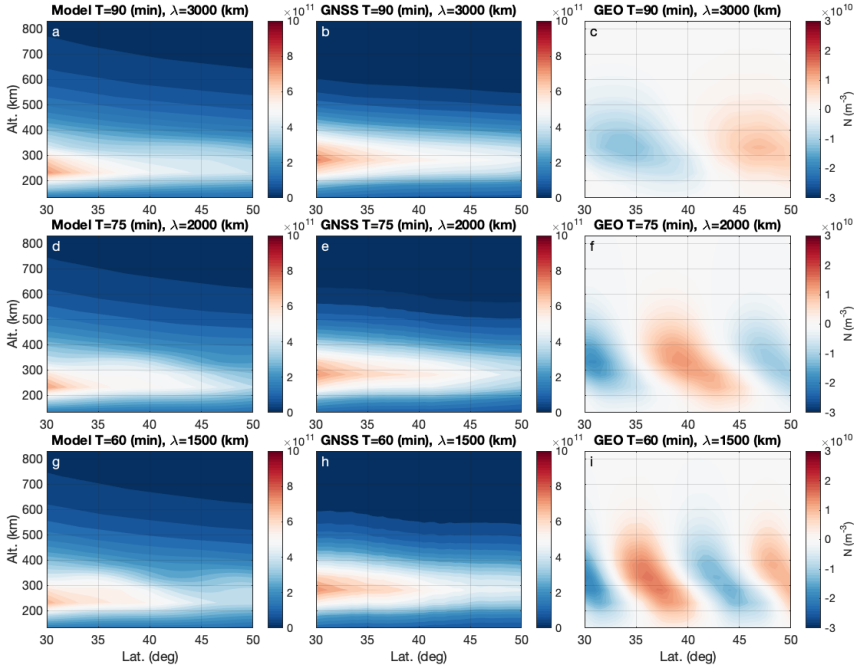

Figure 35: modelled N_{sim} and inversion results N_{est} (m^{-1}) for TIDs 1-3 using ray-differenced TEC. b, e and h show GNSS-MEO results, and c, f and i show GEO-only results.

The results from the ray-differenced GNSS-MEO inversions are similar to the absolute inversions, apart from a small shift in overall background vTEC for some of the inversions. These arise from the need to calibrate the absolute values within the reconstruction algorithm. The artefacts seen in the pre-calibrated results are equally apparent in figures 37e-f and 37j-l (TIDs 4-6), and the amplitudes are again occasionally over-estimated when using GNSS.

7.3.5 Comparing GEO-only and GNSS-MEO results

When using calibrated TEC, both GEO and GNSS-MEO data produce inversions with TIDs clearly visible for all but the smallest TID. The largest error in vTEC between the model and the GEO reconstruction is around 0.25 TECu, after the modelled background N_0 has been removed.

The clearest difference between the GEO and GNSS results is apparent in the ray-differenced TEC results. Unlike the GNSS case, the GEO inversion is essentially insensitive to the slowly changing background N_0, and therefore well suited to image

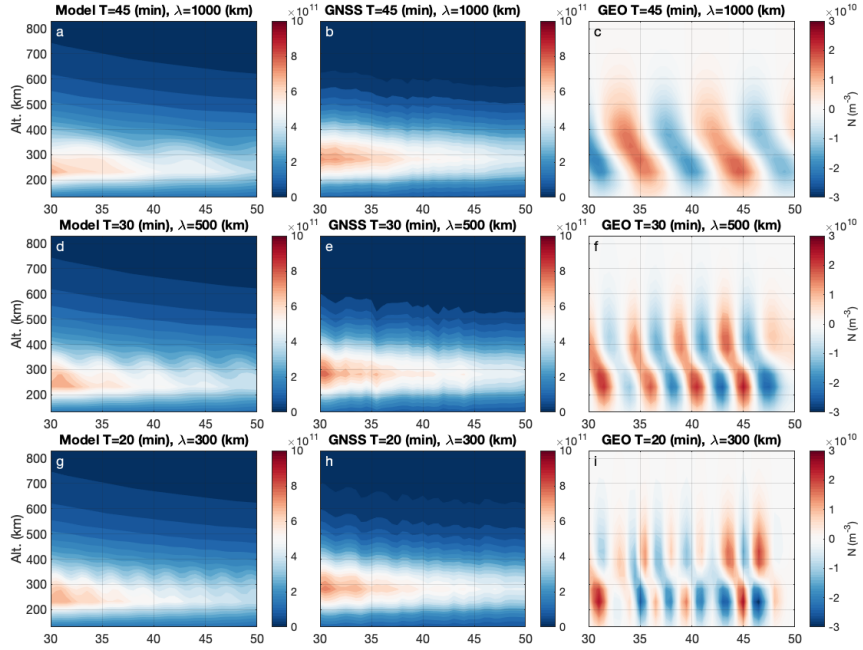

Figure 36: modelled N_{sim} and inversion results N_{est} (m^{-1}) for TIDs 4-6 using ray-differenced TEC. b, e and h show GNSS-MEO results, and c, f and i show GEO-only results.

MSTIDs which more often induce changes in TEC between sequential rays. It is also poorer in reconstructing the larger LSTIDs 1-3 for the same reason. These results indicate that the two satellite geometries are complementary. As the GEO rays are stationary, changes in TEC over 30 s caused by the temporal gradient in N_0 (IRI) are very small in comparison to the changes caused by a TID. Therefore, the GEO-only inversion essentially has no information on the TEC contribution from N_0, and solves only for the more rapid perturbations. For moving satellites, on the other hand, the change in TEC over 30 s is more substantial, as it is also affected by the spatial gradient in N_0.

7.3.6 Comparing LSTID and MSTID results

The results show that the scale of the modelled TID has a large impact on the reconstruction results. The smallest TIDs (5-6) are especially difficult to reconstruct using GNSS-MEO geometry, with exaggerated amplitudes and noisy spatial artefacts in vTEC present in figures 34 and 37. It is likely that these are caused by the distortion effects

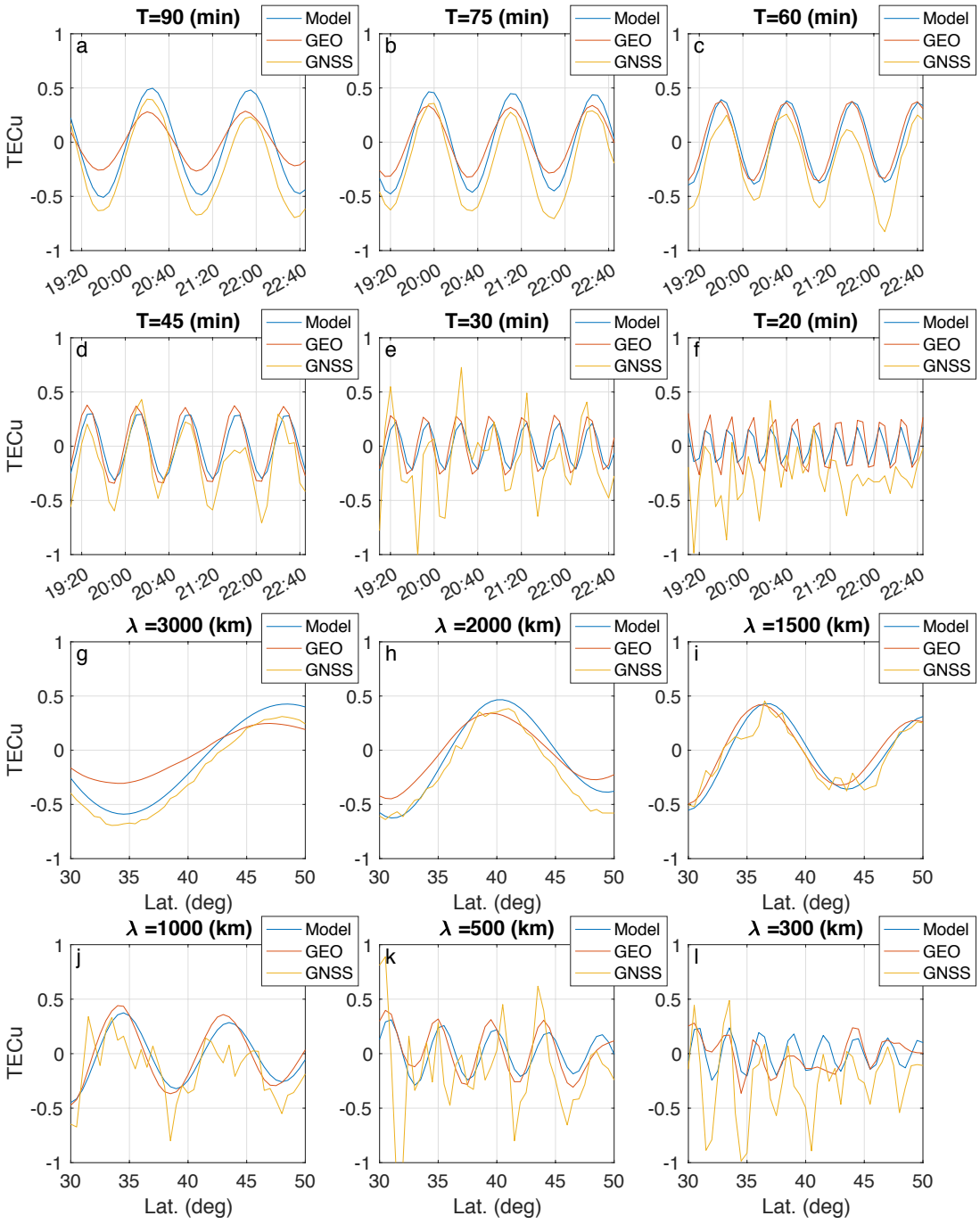

Figure 37: Ray-differenced inversion results ($vTEC_{est}^{GNSS-MEO} - vTEC_0$) and $vTEC_{est}^{GEO}$ plotted against the model δTEC.

due to satellite motion detailed in section 7.1, which become prominent when the data time window of 1 h is comparable to the TID wave period, and produces seemingly inconsistent TEC data. These distortions are not present in the GEO data, making it better suited for imaging MSTIDs.

At the other end of the spectrum, large LSTIDs are imaged slightly better using GNSS than GEO for pre-calibrated TEC inversions. This could be because GNSS simply has a higher number of visible satellites, or due to the limited viewing angles of the GEO satellites, which may introduce some latitudinal dependence in the wave amplitude. GNSS is also better for cases using ray-differenced TEC since the GEO ray-differenced inversion has difficulty detecting these large structures.

It is clear from figures 35, 36, 32 and 33 that all inversions overestimate the peak height of the model. This is a result of the pre-defined Chapman profiles used to seed the EOFs, which provide the vertical structure for the inversion. These are the same for all inversions, and consequently the F-layer peak is consistently overestimated.

7.3.7 Combining GEO and GNSS-MEO geometry

Having now established that the two types of orbits have different effects on the resulting inversion, the next step is to examine the performance of both GEO and GNSS-MEO data together. The results in the above sections suggest that it could be beneficial to combine the GNSS and GEO in order to have a single method to more reliably detect TIDs of all scales. The simplest way to combine both satellite geometries is to use both the GEO and GNSS data sets simultaneously in one inversion without any weighting scheme.

The combined reconstruction is very similar to the GNSS-only reconstruction, as seen in Figure 38. A reason for this could be that there are generally more GNSS data available than GEO data, and that the moving GNSS rays can cover areas of the inversion grid that the stationary GEO rays do not, giving the GNSS data precedence. Essentially, the benefit of using the GEO signals is not realised when combined together with the GNSS data in this way.

7.3.8 Summary of section 7.3

The results of this section have shown that the MIDAS tomography algorithm is able to reconstruct TID-perturbed ionospheres using standard GNSS satellite geometry and GEO geometry. When used separately, there were distinct advantages seen with the different satellite geometries. Tables 4 and 5 summarise the results as root mean square

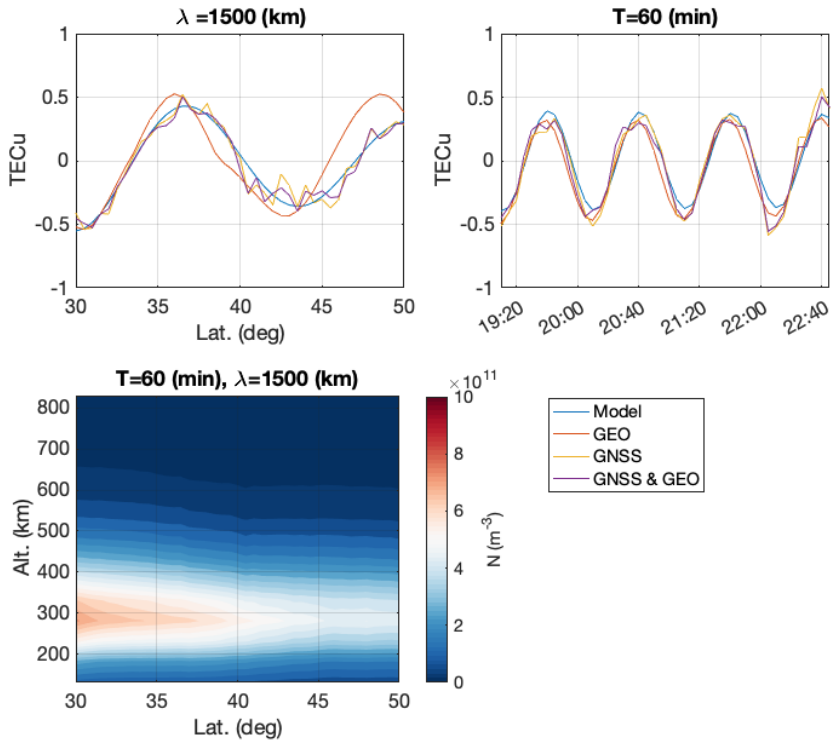

Figure 38: Inversion using pre-calibrated GEO and GNSS-MEO data combined for TID 3.

errors (RMSE) of vTEC sampled at a point in the centre of the grid (40^o N 98^o W) for all sets of TID parameters in Table 3. Table 4 shows the RMSEs using pre-calibrated TEC, which are generally smaller than the errors from ray-differenced MIDAS results in Table 5. In general, both tables show an overall decreasing error with smaller TIDs for GEO-only results, and the opposite trend for GNSS-only and combined results. The combined results have equal or slightly lower RMSE than the GNSS-only inversions for all but the two smallest TIDs, where the distortions from satellite movement are largest.

Table 4: Root mean square errors (TECu) between estimated vertical TEC perturbation ($\delta vTEC_{est}$ for GEO, $\delta vTEC_{est} - vTEC_{IRI}$ for GNSS-MEO and combined) from calibrated TEC inversions and perturbed vTEC of the input model ($\delta vTEC_{sim}$). These correspond to a-f in Figure 34.

Pre-calib.	TID_1	TID_2	TID_3	TID_4	TID_5	TID_6
T (min)	90	75	60	45	30	20
λ (km)	3000	2000	1500	1000	500	300
GEO	0.09	0.09	0.10	0.10	0.10	0.16
GNSS	0.04	0.06	0.10	0.15	0.27	0.18
Combined	0.03	0.06	0.09	0.15	0.26	0.21

Table 5: Root mean square errors (TECu) between estimated vertical TEC perturbation ($\delta vTEC_{est}$ for GEO, $\delta vTEC_{est} - vTEC_{IRI}$ for GNSS-MEO and combined) from ray-differenced TEC inversions and perturbed vTEC of the input model ($\delta vTEC_{sim}$). These correspond to a-f in Figure 37.

Ray-diff.	TID_1	TID_2	TID_3	TID_4	TID_5	TID_6
T (min)	90	75	60	45	30	20
λ (km)	3000	2000	1500	1000	500	300
GEO	0.18	0.12	0.07	0.10	0.10	0.13
GNSS	0.20	0.20	0.23	0.26	0.29	0.31
Combined	0.17	0.19	0.20	0.25	0.32	0.30

7.4 Discussion and conclusions

In this paper, we have used simulated TEC data-sets to investigate the performance of the MIDAS ionospheric tomography algorithm in relation to TID imaging. These data-sets are generated from different stationary and moving satellite geometries through modelled TIDs. The results suggest that GNSS-MEO satellite geometry is preferred for LSTIDs with wave periods above one hour. GEO geometry performs better for medium-scale TIDs.

For GNSS-MEO, there is a clear effect of satellite movement introducing distortions of the TID perturbed TEC, as shown by Wan et al. (1997), and this can have a subsequent effect on the tomographic images produced by MIDAS by distorting the input TEC data. This effect appears to be most prominent for MSTIDs.

When using MIDAS in ray-differenced mode with GEO TEC, the TID structures are reproduced separated from the background ionosphere, even for medium-scale TIDs. Using MIDAS in pre-calibrated mode with GEO TEC produced full inversions including the background ionosphere, and showed improvement over GNSS-only TEC for MSTIDs.

This study did not find that inversion using a direct combination of both GNSS and GEO data offered any major advantages when imaging TIDs. A possible future development of the reconstruction method could use a weighting scheme giving GEO observations precedence during the inversion. This could improve imaging of MSTIDs, while still making use of the full GNSS geometry.

An inherent difference between the GEO and GNSS geometries is the limited viewing angle of a GEO satellite, all of which are positioned over the equator. If the geometry is such that the ray does not cut through the trough and crest of the wave, it can even be the case that the ray intersects two or more wavefronts. In this case an increase in

density where the ray intersects one wavefront is compensated for by a depletion where the ray intersects another wavefront. This is described as the *phase-cancellation effect* by Georges and Hooke (1970). For this reason, there are only a certain range of latitudes where the viewing geometry from a ground-based receiver will be favourable for GEO TID observations.

This study used purely simulated observations from a modelled ionosphere. It still remains to apply these methods on real data known to contain TIDs and to verify against independent observations. When using real GNSS signal observations, there are several considerations to be taken into account. Firstly, the real ionosphere will of course differ from the models used here, and will often be much more dynamic, and not as well behaved. For example, in addition to TIDs there may be other large-scale perturbations present that could obscure the TID signatures. There may also be several TIDs present at once, and the method presented here does not include a strategy for separating them out. The Hooke TID model used here considers TIDs as relative perturbations, so the wave amplitudes scale with the local level of ionisation. This means that for these results, a stronger or weaker ionosphere would not make the TIDs more or less difficult to image. In a real case however, a weaker ionosphere makes the TID TEC perturbations smaller in absolute terms, and so are more difficult to discern in the presence of noise. Secondly, if using un-calibrated real TEC with MIDAS, the MIDAS calibration accuracy can be expected at around 1 TECu (Dear, 2007) and could thus limit reliable detection of TIDs to those with amplitudes larger than 1 TECu. Achieving GNSS IFB calibration to a level better than 1 TECu is challenging, as equipment limitations (temperature variations) are fundamental to the calibration and would have to be modelled into the process to increase the accuracy of the instantaneous IFB estimation. Finally, the number and type of geostationary navigation satellites in view varies geographically. Here, it is assumed that all GNSS receivers receive the signals from three SBAS GEO satellites. In practice some SBAS satellites transmit only one frequency, thus requiring an additional step of processing as described by Cooper et al. (2019). Geostationary satellites that broadcast a full GNSS signal, such as BeiDou GEO satellites, can be used for direct TEC measurements (Kunitsyn et al., 2016) in areas where they provide coverage.

The methods presented in this paper can be applied to GNSS observations over different regions of the world to investigate the spatial and temporal characteristics of TIDs from a statistical viewpoint. These can then be related to proposed mechanisms of TID generation to uncover which mechanisms are prevalent for different TID observations and classes. Interesting outstanding questions about TIDs include the source mechanisms such as thunderstorms and other meteorological disturbances (Azeem and Barlage, 2018),

geomagnetic storms and auroral energy deposition. Other more unusual anthropogenic causes could also be considered, such as those investigated by Scott and Major (2018). These types of investigations require TID tracking over large areas to infer not only the spatial and temporal wave characteristics locally, but also the TID direction and hence the source region. The potential to use distributed GNSS and GEO observations for these types of ionospheric studies is very promising.

Taking into account the challenges that still remain, the geostationary orbit offers a unique vantage point for observing travelling ionospheric disturbances, where the motion of the satellite does not produce frequency shifts to oscillations in TEC, and will provide a valuable input into ionospheric tomography and data assimilation algorithms concerned with TID imaging in the future.

References

Andreeva, E., V. Kunitsyn and E. Tereshchenko (1992). 'Phase-Difference Radio Tomography of the Ionosphere'. In: *Geomagnetism and Aeronomy/Geomagnetizm i Aeronomiia* 32.1, pp. 104–110.

Austen, J. R., S. J. Franke and C. H. Liu (1988). 'Ionospheric imaging using computerized tomography'. In: *Radio Science* 23.3, pp. 299–307. DOI: 10.1029/RS023i003p00299.

Azeem, Irfan and Michael Barlage (2018). 'Atmosphere-ionosphere coupling from convectively generated gravity waves'. In: *Advances in Space Research* 61.7, pp. 1931–1941. DOI: 10.1016/j.asr.2017.09.029.

Bilitza, D. et al. (2017). 'International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions'. In: *Space Weather* 15.2, pp. 418–429. DOI: 10.1002/2016SW001593.

Bruno, Jon et al. (2020). 'A realistic simulation framework to evaluate ionospheric tomography'. In: *Advances in Space Research* 65.3, pp. 891–901. DOI: 10.1016/j.asr.2019.11.015.

Bust, Gary S. and Cathryn N. Mitchell (2008). 'History , Current State, and Future Directions of Ionospheric Imaging'. In: *Reviews of Geophysics* 46.2006, pp. 1–23. DOI: 10.1029/2006RG000212.

Cervera, M. A. and T. J. Harris (2014). 'Modeling ionospheric disturbance features in quasi-vertically incident ionograms using 3-D magnetoionic ray tracing and atmospheric gravity waves'. In: *Journal of Geophysical Research: Space Physics* 119.1, pp. 431–440. DOI: 10.1002/2013JA019247.

Chapman, S. (1931). 'The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth'. In: *Proceedings of the Physical Society*, pp. 26–45.

Chen, C. H. et al. (2016). 'Medium-scale traveling ionospheric disturbances by three-dimensional ionospheric GPS tomography 3. Space science'. In: *Earth, Planets and Space* 68.1. DOI: 10.1186/s40623-016-0412-6.

Cook, J A and S Close (1995). 'An investigation of TID evolution observed in MACE'93 data'. In: *Ann. Geophys.* 13, pp. 1320–1324.

Cooper, C. et al. (2019). 'Measurement of Ionospheric Total Electron Content Using Single-Frequency Geostationary Satellite Observations'. In: *Radio Science* 54.1, pp. 10–19. DOI: 10.1029/2018RS006575.

Davis, M. J. and A. V. da Rosa (1969). 'Traveling ionospheric disturbances originating in the auroral oval during polar substorms'. In: *Journal of Geophysical Research* 74.24, pp. 5721–5735. DOI: 10.1029/JA074i024p05721.

Dear, Richard M. (2007). 'Wide-Area Forecasting of Total Electron Content over Europe'. PhD thesis. University of Bath.

Dear, Richard M. and Cathryn N. Mitchell (2006). 'GPS interfrequency biases and total electron content errors in ionospheric imaging over Europe'. In: *Radio Science* 41.6, pp. 1–7. DOI: 10.1029/2005RS003269.

Georges, T. M. and W. H. Hooke (1970). 'Wave-induced fluctuations in ionospheric electron content: A model indicating some observational biases'. In: *Journal of Geophysical Research* 75.31, pp. 6295–6308. DOI: 10.1029/JA075i031p06295.

Hernández-Pajares, Manuel, J Miguel Juan and Jaume Sanz (2006a). 'Medium-scale traveling ionospheric disturbances affecting GPS measurements: Spatial and temporal analysis'. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–13. DOI: 10.1029/2005JA011474.

– (2006b). 'Real Time MSTIDs Modelling and Application to Improve the Precise GPS and GALILEO Navigation'. In: *Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006)*. Fort Worth, Texas, pp. 1358–1368.

Hines, C. O. (1960). 'Internal Amospheric Gravity Waves at Ionospheric Heights'. In: *Canadian Journal of Physics* 38.11, pp. 1441–1481. DOI: 10.1139/p60-150.

Hocke, K. and K. Schlegel (1996). 'A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995'. In: *Annales Geophysicae* 14.9, pp. 917–940. DOI: 10.1007/s00585-996-0917-6.

Hooke, W. H. (1968). 'Ionospheric irregularities produced by internal atmospheric gravity waves'. In: *Journal of Atmospheric and Terrestrial Physics* 30.5, pp. 795–823. DOI: 10.1016/S0021-9169(68)80033-9.

Huang, Fuqing et al. (2018). 'Nighttime Medium-Scale Traveling Ionospheric Disturbances From Airglow Imager and Global Navigation Satellite Systems Observations'. In: *Geophysical Research Letters* 45.1, pp. 31–38. DOI: 10.1002/2017GL076408.

Kunitake, M et al. (1995). 'Tomographic imaging of the ionosphere over Japan by the modified truncated SVD method'. In: *Ann. Geophys.* 13, pp. 1303–1310.

Kunitsyn, V. E. et al. (2016). 'Ionospheric TEC estimation with the signals of various geostationary navigational satellites'. In: *GPS Solutions* 20.4, pp. 877–884. DOI: 10.1007/s10291-015-0500-2.

Lee, Jeffrey K., Farzad Kamalabadi and Jonathan J. Makela (2008). 'Three-dimensional tomography of ionospheric variability using a dense GPS receiver array'. In: *Radio Science* 43.3, pp. 1–15. DOI: 10.1029/2007RS003716.

Mitchell, C. N., N. R. Rankov et al. (July 2017). 'Ionospheric data assimilation applied to HF geolocation in the presence of traveling ionospheric disturbances'. In: *Radio Science* 52.7, pp. 829–840. DOI: 10.1002/2016RS006187.

Mitchell, Cathryn N. and Paul S.J. Spencer (2003). 'A three-dimensional time-dependent algorithm for ionospheric imaging using GPS'. In: *Annals of Geophysics* 46.4, pp. 687–696. DOI: 10.4401/ag-4373.

Morgan, M. G. and C. H. J. Calderón (1978). 'Testing Hooke's Perturbation Formula for the Production of TID's by Gravity Waves'. In: *Journal of Geophysical Research* 83.A12, pp. 5737–5740.

Munro, G. H. (July 1950). 'Travelling disturbances in the ionosphere'. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 202.1069, pp. 208–223. DOI: 10.1098/rspa.1950.0095.

Nickisch, L. J. et al. (2016). 'Feasibility study for reconstructing the spatial-temporal structure of TIDs from high-resolution backscatter ionograms'. In: *Radio Science* 51.5, pp. 443–453. DOI: 10.1002/2015RS005906.

Nygrén, T et al. (May 2015). 'Radar observations of simultaneous traveling ionospheric disturbances and atmospheric gravity waves'. In: *Journal of Geophysical Research: Space Physics* 120.5, pp. 3949–3960. DOI: 10.1002/2014JA020794.

Penney, R. W. and N. K. Jackson-Booth (2015). 'Mitigating satellite motion in GPS monitoring of traveling ionospheric disturbances'. In: *Radio Science* 50.11, pp. 1150–1164. DOI: 10.1002/2015RS005767.

Pryse, S E et al. (1995). 'Travelling ionospheric disturbances imaged by tomographic techniques'. In: *Annales geophysicae* 13.12, pp. 1325–1330.

Saito, A., S. Fukao and S. Miyazaki (1998). 'High resolution mapping of TEC perturbations with the GSI GPS network over Japan'. In: *Geophysical Research Letters* 25.16, pp. 3079–3082. DOI: 10.1029/98GL52361.

Scott, Christopher J. and Patrick Major (2018). 'The ionospheric response over the UK to major bombing raids during World War II'. In: *Annales Geophysicae* 36.5, pp. 1243–1254. DOI: 10.5194/angeo-36-1243-2018.

Ssessanga, Nicholas, Yong Ha Kim and Eunsol Kim (2015). 'Vertical structure of medium-scale traveling ionospheric disturbances'. In: *Geophysical Research Letters* 42.21, pp. 9156–9165. DOI: 10.1002/2015GL066093.

Tang, Jun et al. (2016). 'Large-scale traveling ionospheric disturbances using ionospheric imaging at storm time: A case study on 17 march 2013'. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 145, pp. 12–20. DOI: 10.1016/j.jastp.2016.04.006.

Thébault, Erwan et al. (2015). 'International geomagnetic reference field: The 12th generation international geomagnetic reference field - The twelfth generation'. In: *Earth, Planets and Space* 67.1. DOI: 10.1186/s40623-015-0228-9.

Van De Kamp, M., D. Pokhotelov and K. Kauristie (2014). 'TID characterised using joint effort of incoherent scatter radar and GPS'. In: *Annales Geophysicae* 32.12, pp. 1511–1532. DOI: 10.5194/angeo-32-1511-2014.

Wan, Weixing et al. (1997). 'TID Observation Using a Short Baseline Network of Gps Receivers'. In: *Acta Geodaetica et Geophysica Hungarica* 32.3-4, pp. 321–327. DOI: 10.1007/BF03325503.

Yin, Ping et al. (2017). 'A multiresolution inversion for imaging the ionosphere'. In: *Journal of Geophysical Research: Space Physics* 122.6, pp. 6799–6811. DOI: 10.1002/2016JA023728.

Yizengaw, Ekassie et al. (2005). 'Ionosphere dynamics over the Southern Hemisphere during the 31 March 2001 severe magnetic storm using multi-instrument measurement data'. In: *Annales Geophysicae* 23.3, pp. 707–721. DOI: 10.5194/angeo-23-707-2005.

Part III

8 | Conclusions and outlook

Summary of conclusions

Part I introduced the fundamentals of GNSS, the ionosphere, ionospheric tomography and TID monitoring. This aimed to give the necessary theoretical background, and to illustrate the importance of monitoring TIDs in order to mitigate their impact on precise real-time kinematic GNSS positioning and to understand more about their origin and behaviour. To tackle these problems, MIDAS, a well-established tool for ionospheric data assimilation and tomography, was also described in this part of the thesis.

In part II, Chapter 5 first quantified the impact that TIDs can have on network-generated ionospheric corrections for N-RTK, and showed that TID-induced errors can be large enough to pose a problem, especially for sparse networks. These errors appeared in the interpolation of the ionospheric corrections, and several interpolation methods were therefore tested and compared. As a potential solution, interpolation techniques were also modified to include TID information in order to better interpolate with TIDs present. This required TID information to be extracted from measurements as a first step. Several methods to extract TID information from GNSS that work directly with TEC time-series have been developed by other researchers, and one such method (Hernández-Pajares, Juan and Sanz, 2006) was applied in this chapter. Using the estimated TID direction and wavelength to change the weighting scheme for a WLS low-order surface interpolation was shown to decrease the number of large interpolation errors during the passage of TIDs.

An alternative approach to TID mitigation is to use ionospheric electron density maps, such as those generated by ionospheric tomography. TID imaging using the MIDAS tomography algorithm was therefore tested under different circumstances, the first being a case study of a TID occurring on 31 October 2003 during the recovery phase of the 2003 Halloween Storm. MIDAS was run with data from a relatively sparse North-American

network with only GPS available. The resulting tomographic images showed a large-amplitude TID propagating towards the equator, as is typical for LSTIDs launched in the auroral oval during geomagnetic storms.The results were compared with ionosonde soundings and in situ measurements from the LEO CHAMP satellite. The comparisons showed that in the regions with the poorest receiver coverage, perturbations from the TID were missing in the MIDAS results. However, it was shown that a modern network would have shown the TID accurately in the same areas.

Chapter 7 builds on the initial results to examine to which extent MIDAS is able to image TIDs of different scales. The modern network used for the simulation in Chapter 6 was now used with simulated data from GPS, GLONASS and three satellites in GEO. In these scenarios, it is concluded that large-scale TIDs are well reconstructed using GPS and GLONASS, and that the imaging of medium-scale TIDs is improved by using geostationary orbits. A purely stationary geometry was also shown to have the added benefit of isolating the perturbations in the images. If the tomography simulation results are compared to the interpolation simulation results from Chapter 5, it is evident that for λ/b ratios above 2 the MIDAS pre-calibrated results perform on par with most interpolation techniques in terms of TID-induced errors. The RMSE of the modelled TIDs in Tables 4 and 5 from Chapter 7 are plotted in the style of Figure 11 in Figure 39.

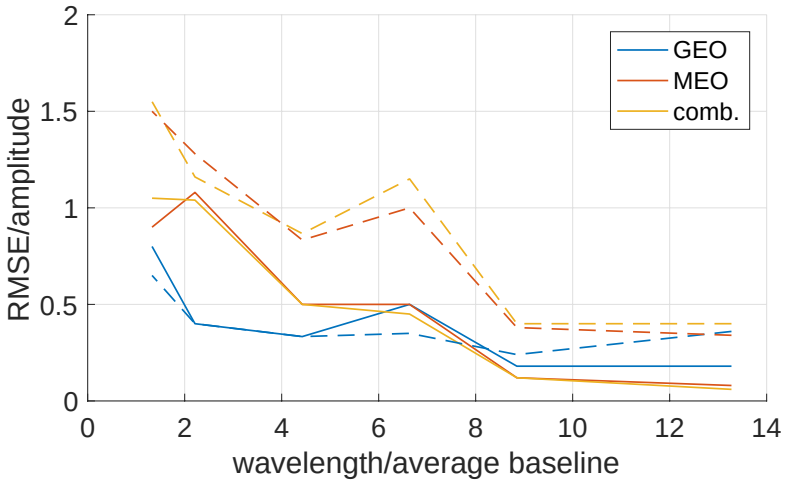

Figure 39: RMSE of the simulated tomography results in Chapter 7. Filled lines represent pre-calibrated mode and dashed lines ray-differenced mode.

Outlook

The central role of receiver geometry when observing TIDs is a recurring theme throughout part II, and the continued construction of dense networks throughout the world makes TID observation using GNSS increasingly effective. The wavelength-network baseline ratio, a proxy for receiver density, was shown in Chapter 5 to be the most important factor affecting the severity of the interpolation error caused by simulated TIDs. The modified interpolation techniques tackled this problem by introducing a variable weighting, allocating higher weights to nearby receivers when the distances between receivers is large in comparison to the TID wavelength. As expected, the network density also had a profound effect on the MIDAS reconstructions. The network used in the simulations of Chapter 7 was sufficient to image simulated waves, while the LSTID imaged in Chapter 6 was reasonably well reconstructed over the denser parts of the network, poor performance was noted in areas with few receivers. An interesting technique to address this issue has been used by e.g. S. Saito et al. (2017), who used large voxels ($5° \times 5°$) at the edges of the network, and progressively smaller voxels towards the denser centre of the network ($2° \times 2°$ and $1° \times 1°$). The implementation of adaptive, receiver-density dependent voxel size would therefore be a possible next step in the development of MIDAS.

One key to improving not only the capabilities of MIDAS, but also that of many other ionospheric monitoring techniques is the incorporation of TEC observations from geostationary orbit, as exemplified in Chapter 7. Geostationary satellites have the benefit over MEO satellites that there are no movement-induced distortions to the apparent TID period, which allows for a more straightforward interpretation of the data. Kunitsyn et al. (2016) measured TEC from available multi-frequency GEO satellites in the BeiDou GNSS system and the European (EGNOS), American (WAAS) and Indian (GAGAN) SBAS systems. They concluded that the BeiDou noise levels in the TEC estimated from the BeiDou satellites were one order smaller than those of the SBAS systems. The BeiDou GEO satellites, however, do not have global coverage as they are located between 58.5°E and 160°E. The possible addition of satellites in GEO that can enable TEC estimates with GNSS-level accuracy to Galileo, GPS and GLONASS are therefore an important consideration, as these can cover over other parts of the world. An alternative is to use single-frequency signals, as done by Cooper et al. (2019) who introduced a method to estimate TEC from single-frequency signals from GEO satellites. They were able to get reliable TEC estimations from the EGNOS SES-5 satellite with three European ground receivers. To enable widespread use of this technique more high-quality receivers need to be made available, since not all receivers are able to

provide quality measurements using this technique.

Currently, BeiDou provides the best available combination of available ground receiver networks and satellites for TID studies using GEO TEC. For example, reference networks in China and Australia provide interesting opportunities for larger statistical studies of TIDs using GEO-TEC, using the tomographic methods of Chapter 7 or frequency-domain methods like those of Katamzi et al. (2012). Such studies would offer an opportunity to study TIDs and other ionospheric perturbations and to improve the understanding of TID generation and climatology throughout the region. Ionospheric perturbations in BeiDou GEO TEC from a receiver in the Auscope reference network in Australia is plotted in Figure 40, as an example of what data is available. Such data from a dense network could potentially be used with MIDAS to image the perturbations in isolation, following the work in Chapter 7.

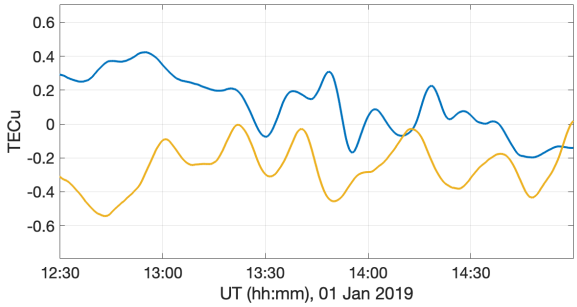

Figure 40: Ionospheric perturbations in un-calibrated TEC from two BeiDou GEO satellites seen from the Auscope Arubiddy reference station (31.8°S, 125.9°E).

Interesting avenues for future research can also be found in development of the TID imaging and mitigation algorithms themselves. Although they are presented as alternative methods in this thesis, the interpolation techniques of Chapter 5 and the tomography approach used in chapters 6 –7 could potentially be combined in two interesting ways. Firstly, the 3D tomographic images of TIDs presented in this work could potentially be used to automatically estimate TID wave parameters. This could technically be done from by applying 3D spectral methods, e.g. the Stockwell transform (Stockwell et al., 1996) which was successfully applied to detect and characterise gravity waves in satellite data from the Atmospheric Infrared Sounder (AIRS) by Hindley et al. (2019). If this were implemented, it could potentially be used in place of the TEC time-series based method used in Chapter 5. Secondly, instead of interpolating between the reference receiver locations, as done in Chapter 5, it is possible to use the same interpolation techniques between individual pixels in a vTEC map. The results in chapters 7-6 have

shown that TIDs can be reconstructed by MIDAS, so a tomography-generated vTEC map could be sampled more accurately for the approximate user position.

This thesis has shown that MIDAS in its current configuration can be used directly to image LSTIDs and some MSTIDs. This also means that their effect on positioning using ionospheric corrections based on tomographic electron density maps can potentially include TIDs, and therefore mitigate potential effects on high-accuracy positioning. The use of denser networks, multi-GNSS and, in particular, geostationary measurements could make the tomographic imaging of smaller MSTIDs more reliable. When this is not available, the modified interpolation methods tested in this thesis are an alternative solution. However, these require a reliable estimation of TID parameters which adds an additional processing step, which is also affected by receiver-satellite geometry.

In practice, the TID-induced errors studied in this thesis are only part of the total ionospheric delay, and the remaining differences between the unperturbed true ionosphere and the tomographic models or images may be even larger. A wealth of information on the performance of MIDAS with this aspect in focus can be found in (Bruno, 2020). Fortunately, the ways to improve TID imaging using tomography, i.e. improved receiver-satellite geometry, are also ways to improve the imaging of the background ionosphere. This makes tomographic imaging an attractive solution for future reliable high-accuracy GNSS positioning solutions.

List of acronyms

TID	Travelling Ionospheric Disturbance
LSTID	Large-Scale TID
MSTID	Medium-Scale TID
SSTID	Small-Scale TID
AGW	Atmospheric Gravity Wave
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
BDS	BeiDou Navigation Satellite System
GLONASS	GLObalnaya NAvigatsionnaya Sputnikovaya Sistema, or GLObal NAvigation Satellite System
EUV	Extreme Ultra-Violet
CME	Coronal Mass Ejection
TEC	Total Electron Content
vTEC	vertical TEC
sTEC	slant TEC
IPP	Ionospheric Pierce Point
MEO	Medium Earth Orbit
GEO	Geostationary Earth Orbit
LEO	Low Earth Orbit
GSO	Geosynchronous Orbit
IGSO	Inclined Geosynchronous Orbit
PRN	Pseudo-Random Noise
SBAS	Satellite-Based Augmentation System
GBAS	Ground-Based Augmentation System
PPP	Precise Point Positioning

RTK	Real-Time Kinematic
N-RTK	Network-RTK
PPP	Precise Point Positioning
CIT	Computerised Ionospheric Tomography
RT	Radio Tomography
SVD	Singular Value Decomposition
TSVD	Truncated SVD
MIDAS	Multi-Instrument Data Analysis System
IFB	Inter-Frequency Bias
EOF	Empirical Orthogonal Function
ISR	Incoherent Scatter Radar
ASC	All-Sky Camera
GEONET	GPS Earth Observation Network
FFT	Fast Fourier Transform
MART	Multiplicative ALgebraic Reconstruction Algorithm
IRI	International Reference Ionosphere
IGRF	International Geomagnetic Reference Field
IGS	International GNSS Service
WAAS	Wide Area Augmentation System
RMSE	Root Mean Square Error
PLP	Planar Langmuir Probe
UTC	Coordinated Universal Time
MLT	Magnetic Local Time
AE	Auroral Electrojet
SSR	State Space Representation
AR	Ambiguity Resolution
IDW	Inverse Distance Weighting
OK	Ordinary Kriging
WLS	Weighted Least Squares
WLSD	Weighted Least Squares with Direction
WLSL	Weighted Least Squares with waveLength
NETPOS	NEtherlands POsitioning Service
LGLN	Landesamt für Geoinformation Niedersachsen

List of symbols

λ_D	Debye length
N_e	Electron number density
N_i	Ion number density
ϵ_0	Permittivity of free space
k_B	Boltzmann constant
T_e	Electron temperature
q_e	Electron (elementary) charge
hf	Photon
e^-	Electron
X	Neutral particle species
Y	Neutral particle species
ω	Angular frequency
ω_a	Acoustic cut-off angular frequency
ω_b	Buoyancy angular frequency
$\mathbf{k}$	Wave vector
s	Sound velocity
γ	Heat capacity ratio
g	Gravitational acceleration
$\mathbf{B}$	Magnetic (field) flux density
$\mathbf{u}$	Particle displacement
λ	Wavelength
T	Wave period
v_ϕ	Phase velocity
v_g	Group velocity
n	Index of refraction
ν	Collision frequency

ω_g	Gyrofrequency
ω_p	Plasma frequency
m_e	Electron mass
f	Frequency
c	Speed of light in vacuum
ρ	Geometric range
P	Pseudorange observable
dt_r	Receiver clock bias
dt_s	Satellite clock bias
δt_{rel}	Combined general and special relativity biases
d_r	Receiver instrument bias
d_s	Satellite instrument bias
I	Ionospheric delay
X	Tropospheric delay
ε	Combined multipath and receiver noise
φ	Carrier-phase observable
N	Integer ambiguity
$\mathbf{A}$	Inverse problem geometry matrix
$\mathbf{x}$	Inverse problem model vector
$\mathbf{b}$	Inverse problem data vector
ϱ	Tikhonov regularisation parameter
$\mathbf{R}$	Regularisation matrix
$\mathbf{V}$	Model-space matrix
$\mathbf{U}$	Data-space matrix
$\mathbf{S}$	Singular value matrix
β	Combined bias and ambiguity vector
$\mathbf{s}$	Slowness vector
U_b	Neutral velocity along geomagnetic field
I_B	Geomagnetic inclination
w	Interpolation weight
d	Interpolation distance
p	Interpolation weighting power
b	Reference network average baseline
θ_{TID}	TID direction (from North)

List of Figures

List of Tables

Bibliography

Altadill, D. et al. (2020). 'Identification and monitoring techniques of TIDs in the H2020 TechTIDE project'. In: *EGU General Assembly 2020*. Online. DOI: 10.5194/egusphere-egu2020-7769.

Austen, J. R., S. J. Franke and C. H. Liu (1988). 'Ionospheric imaging using computerized tomography'. In: *Radio Science* 23.3, pp. 299–307. DOI: 10.1029/RS023i003p00299.

Borovsky, Joseph E (Aug. 2020). 'What magnetospheric and ionospheric researchers should know about the solar wind'. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 204, p. 105271. DOI: 10.1016/j.jastp.2020.105271.

Borries, Claudia et al. (2017). 'On the dynamics of large-scale traveling ionospheric disturbances over Europe on 20 November 2003'. In: *Journal of Geophysical Research: Space Physics* 122.1, pp. 1199–1211. DOI: 10.1002/2016JA023050.

Breitling, W. J., R. A. Kupferman and G. J. Gassmann (1967). 'Traveling ionospheric disturbances associated with nuclear detonations'. In: *Journal of Geophysical Research* 72.1, p. 307. DOI: 10.1029/JZ072i001p00307.

Bruno, Jon (2020). 'Ionospheric tomography and data assimilation'. PhD thesis. University of Bath.

Bruno, Jon et al. (2020). 'A realistic simulation framework to evaluate ionospheric tomography'. In: *Advances in Space Research* 65.3, pp. 891–901. DOI: 10.1016/j.asr.2019.11.015.

Bust, Gary S. and Cathryn N. Mitchell (2008). 'History , Current State, and Future Directions of Ionospheric Imaging'. In: *Reviews of Geophysics* 46.2006, pp. 1–23. DOI: 10.1029/2006RG000212.

Chapman, S. (1931). 'The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth'. In: *Proceedings of the Physical Society*, pp. 26–45.

Chartier, Alex T. et al. (2012). 'The use of ionosondes in GPS ionospheric tomography at low latitudes'. In: *Journal of Geophysical Research: Space Physics* 117.10, pp. 1–9. DOI: 10.1029/2012JA018054.

Chen, C. H. et al. (2016). 'Medium-scale traveling ionospheric disturbances by three-dimensional ionospheric GPS tomography 3. Space science'. In: *Earth, Planets and Space* 68.1. DOI: 10.1186/s40623-016-0412-6.

Cooper, C. et al. (2019). 'Measurement of Ionospheric Total Electron Content Using Single-Frequency Geostationary Satellite Observations'. In: *Radio Science* 54.1, pp. 10–19. DOI: 10.1029/2018RS006575.

Davis, M. J. and A. V. Da Rosa (1970). 'Possible detection of atmospheric gravity waves generated by the solar eclipse'. In: *Nature* 226.5251, p. 1123. DOI: 10.1038/2261123a0.

Davis, M. J. and A. V. da Rosa (1969). 'Traveling ionospheric disturbances originating in the auroral oval during polar substorms'. In: *Journal of Geophysical Research* 74.24, pp. 5721–5735. DOI: 10.1029/JA074i024p05721.

Dear, Richard M. and Cathryn N. Mitchell (2006). 'GPS interfrequency biases and total electron content errors in ionospheric imaging over Europe'. In: *Radio Science* 41.6, pp. 1–7. DOI: 10.1029/2005RS003269.

Garcia, Rolando R. et al. (2017). 'Modification of the gravity wave parameterization in the Whole Atmosphere Community Climate Model: Motivation and results'. In: *Journal of the Atmospheric Sciences* 74.1, pp. 275–291. DOI: 10.1175/JAS-D-16-0104.1.

Gordon, W. (Nov. 1958). 'Incoherent Scattering of Radio Waves by Free Electrons with Applications to Space Exploration by Radar'. In: *Proceedings of the IRE* 46.11, pp. 1824–1829. DOI: 10.1109/JRPROC.1958.286852.

Habarulema, John Bosco et al. (Jan. 2018). 'Storm Time Global Observations of Large-Scale TIDs From Ground-Based and In Situ Satellite Measurements'. In: *Journal of Geophysical Research: Space Physics* 123.1, pp. 711–724. DOI: 10.1002/2017JA024510.

Hajkowicz, L.A. and R.D. Hunsucker (June 1987). 'A simultaneous observation of large-scale periodic TIDs in both hemispheres following an onset of auroral disturbances'. In: *Planetary and Space Science* 35.6, pp. 785–791. DOI: 10.1016/0032-0633(87)90038-9.

Hansen, Per Christian (Dec. 1987). 'The truncated SVD as a method for regularization'. In: *BIT Numerical Mathematics* 27.4, pp. 534–553. DOI: 10.1007/BF01937276.

Hargreaves, J. K. (1992). *The Solar-Terrestrial Environment*. Cambridge University Press. DOI: 10.1017/CBO9780511628924.

Heisler, LH (1958). 'Anomalies in Ionosonde Records Due to Travelling Ionospheric Disturbances'. In: *Australian Journal of Physics* 11.1, p. 79. DOI: 10.1071/PH580079.

Hernández-Pajares, Manuel, J Miguel Juan and Jaume Sanz (2006). 'Medium-scale traveling ionospheric disturbances affecting GPS measurements: Spatial and temporal analysis'. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–13. DOI: 10.1029/2005JA011474.

Hernández-Pajares, Manuel, J. M. Juan, J. Sanz and A. Aragón-Àngel (Dec. 2012). 'Propagation of medium scale traveling ionospheric disturbances at different latitudes and solar cycle conditions'. In: *Radio Science* 47.6. DOI: 10.1029/2011RS004951.

Hindley, Neil P. et al. (Dec. 2019). 'Gravity waves in the winter stratosphere over the Southern Ocean: high-resolution satellite observations and 3-D spectral analysis'. In: *Atmospheric Chemistry and Physics* 19.24, pp. 15377–15414. DOI: 10.5194/acp-19-15377-2019.

Hines, C. O. (1960). 'Internal Amospheric Gravity Waves at Ionospheric Heights'. In: *Canadian Journal of Physics* 38.11, pp. 1441–1481. DOI: 10.1139/p60-150.

– (Apr. 1967). 'On the nature of traveling ionospheric disturbances launched by low-altitude nuclear explosions'. In: *Journal of Geophysical Research* 72.7, pp. 1877–1882. DOI: 10.1029/JZ072i007p01877.

Hocke, K. and K. Schlegel (1996). 'A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995'. In: *Annales Geophysicae* 14.9, pp. 917–940. DOI: 10.1007/s00585-996-0917-6.

Huang, Fuqing et al. (2018). 'Nighttime Medium-Scale Traveling Ionospheric Disturbances From Airglow Imager and Global Navigation Satellite Systems Observations'. In: *Geophysical Research Letters* 45.1, pp. 31–38. DOI: 10.1002/2017GL076408.

Huang, X. et al. (2016). 'Comparing TID simulations using 3-D ray tracing and mirror reflection'. In: *Radio Science* 51.4, pp. 337–343. DOI: 10.1002/2015RS005872.

Hunsucker, Robert D. (1982). 'Atmospheric gravity waves generated in the high-latitude ionosphere: A review'. In: *Reviews of Geophysics* 20.2, p. 293. DOI: 10.1029/RG020i002p00293.

Kamalabadi, Farzad et al. (2002). 'Tomographic studies of aeronomic phenomena using radio and UV techniques'. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 64.12-14, pp. 1573–1580. DOI: 10.1016/S1364-6826(02)00096-2.

Katamzi, Z.T. et al. (Jan. 2012). 'Statistical analysis of travelling ionospheric disturbances using TEC observations from geostationary satellites'. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 74, pp. 64–80. DOI: 10.1016/j.jastp.2011.10.006.

Kherani, E. A. et al. (2016). 'Traveling ionospheric disturbances propagating ahead of the Tohoku-Oki tsunami: A case study'. In: *Geophysical Journal International* 204.2, pp. 1148–1158. DOI: 10.1093/gji/ggv500.

Kirchengast, Gottfried, Klemens Hocke and Kristian Schlegel (Sept. 1995). 'Gravity waves determined by modeling of traveling ionospheric disturbances in incoherent-scatter radar measurements'. In: *Radio Science* 30.5, pp. 1551–1567. DOI: 10.1029/95RS02080.

Kunitsyn, V. E. et al. (2016). 'Ionospheric TEC estimation with the signals of various geostationary navigational satellites'. In: *GPS Solutions* 20.4, pp. 877–884. DOI: 10.1007/s10291-015-0500-2.

Kutiev, Ivan, Pencho Marinov and Anna Belehaki (2016). 'Real time reconstruction of 3-D electron density distribution over Europe with TaD profiler'. In: *Radio Science* 17, pp. 1176–1187. DOI: 10.1002/2015RS005932.

Lerfald, G.M. et al. (Apr. 1972). 'Traveling ionospheric disturbances observed near the time of the solar eclipse of 7 March 1970'. In: *Journal of Atmospheric and Terrestrial Physics* 34.4, pp. 733–741. DOI: 10.1016/0021-9169(72)90161-4.

Mitchell, Cathryn N. and Paul S.J. Spencer (2003). 'A three-dimensional time-dependent algorithm for ionospheric imaging using GPS'. In: *Annals of Geophysics* 46.4, pp. 687–696. DOI: 10.4401/ag-4373.

Munro, G. H. (1948). 'Short-Period Changes in the F Region of the Ionosphere'. In: *Nature* 162, pp. 886–887. DOI: 10.1038/162680a0.

– (July 1950). 'Travelling disturbances in the ionosphere'. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 202.1069, pp. 208–223. DOI: 10.1098/rspa.1950.0095.

Nappo, Carmen J. (2013). *An introduction to atmospheric gravity waves*. 2nd.

Nicolls, Michael J. and Craig J. Heinselman (2007). 'Three-dimensional measurements of traveling ionospheric disturbances with the Poker Flat Incoherent Scatter Radar'. In: *Geophysical Research Letters* 34.21, pp. 1–6. DOI: 10.1029/2007GL031506.

Nicolls, Michael J., Michael C. Kelley et al. (2004). 'Imaging the structure of a large-scale TID using ISR and TEC data'. In: *Geophysical Research Letters* 31.9, pp. 2–5. DOI: 10.1029/2004GL019797.

Noble, S. T. (1990). 'A large-amplitude traveling ionospheric disturbance excited by the space shuttle during launch'. In: *Journal of Geophysical Research* 95.A11, p. 19037. DOI: 10.1029/JA095iA11p19037.

Nygrén, T et al. (May 2015). 'Radar observations of simultaneous traveling ionospheric disturbances and atmospheric gravity waves'. In: *Journal of Geophysical Research: Space Physics* 120.5, pp. 3949–3960. DOI: 10.1002/2014JA020794.

Ogawa, T. et al. (2002). 'Observations and modeling of 630 nm airglow and total electron content associated with traveling ionospheric disturbances over Shigaraki, Japan'. In: *Earth, Planets and Space* 54.1, pp. 45–56.

Prölss, Gerd W. (2004). *Physics of the Earth's space environment*. Springer. DOI: 10.1007/978-3-642-97123-5.

Rajesh, P. K. et al. (2016). 'Space-based imaging of nighttime medium-scale traveling ionospheric disturbances using FORMOSAT-2/ISUAL 630.0 nm airglow observations'. In: *Journal of Geophysical Research: Space Physics* 121.5, pp. 4769–4781. DOI: 10.1002/2015JA022334.

Raymund, T. D. et al. (1990). 'Application of computerized tomography to the investigation of ionospheric structures'. In: *Radio Science* 25.5, pp. 771–789. DOI: 10.1029/RS025i005p00771.

Reinisch, Bodo et al. (2018). 'Pilot Ionosonde Network for Identification of Traveling Ionospheric Disturbances'. In: *Radio Science* 53.3, pp. 365–378. DOI: 10.1002/2017RS006263.

Rius, A., G. Ruffini and L. Cucurull (1997). 'Improving the vertical resolution of ionospheric tomography with GPS occultations'. In: *Geophysical Research Letters* 24.18, pp. 2291–2294. DOI: 10.1029/97GL52283.

Saito, A., S. Fukao and S. Miyazaki (1998). 'High resolution mapping of TEC perturbations with the GSI GPS network over Japan'. In: *Geophysical Research Letters* 25.16, pp. 3079–3082. DOI: 10.1029/98GL52361.

Saito, S. et al. (2017). 'Real-Time Ionosphere Monitoring by Three-Dimensional Tomography over Japan'. In: *Navigation, Journal of the Institute of Navigation* 64.4, pp. 495–504. DOI: 10.1002/navi.213.

Savastano, Giorgio et al. (June 2017). 'Real-Time Detection of Tsunami Ionospheric Disturbances with a Stand-Alone GNSS Receiver: A Preliminary Feasibility Demonstration'. In: *Scientific Reports* 7.1, p. 46607. DOI: 10.1038/srep46607.

Seibert, Günther (2007). *The history of sounding rockets and their contribution to European space research*. Tech. rep. 647 SP, pp. 39–46.

Shults, Ksenia, Elvira Astafyeva and Sévan Adourian (2016). 'Ionospheric detection and localization of volcano eruptions on the example of the April 2015 Calbuco events'. In: *Journal of Geophysical Research: Space Physics* 121.10, pp. 10, 303–10, 315. DOI: 10.1002/2016JA023382.

Ssessanga, Nicholas, Yong Ha Kim and Eunsol Kim (2015). 'Vertical structure of medium-scale traveling ionospheric disturbances'. In: *Geophysical Research Letters* 42.21, pp. 9156–9165. DOI: 10.1002/2015GL066093.

Stefanello, M. B. et al. (June 2015). 'OI 630.0nm all-sky image observations of medium-scale traveling ionospheric disturbances at geomagnetic conjugate points'. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 128, pp. 58–69. DOI: 10.1016/j.jastp.2015.03.012.

Stockwell, R.G., L. Mansinha and R.P. Lowe (Apr. 1996). 'Localization of the complex spectrum: the S transform'. In: *IEEE Transactions on Signal Processing* 44.4, pp. 998–1001. DOI: 10.1109/78.492555.

Teunissen, P. J. G. (1994). 'A new method for fast carrier phase ambiguity estimation'. In: *Proceedings of 1994 IEEE Position, Location and Navigation Symposium - PLANS'94*. Las Vegas, NV, USA, IEEE, pp. 562–573. DOI: 10.1109/PLANS.1994.303362.

Teunissen, Peter J.G. and Oliver Montenbruck (2017). *Springer Handbook of Global Navigation Satellite Systems*. Ed. by Peter J.G. Teunissen and Oliver Montenbruck. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-42928-1.

Thome, George D. (Oct. 1964). 'Incoherent scatter observations of traveling ionospheric disturbances'. In: *Journal of Geophysical Research* 69.19, pp. 4047–4049. DOI: 10.1029/JZ069i019p04047.

Tikhonov, Andrey N and Vasiliy Y Arsenin (1977). *Solutions of ill-posed problems*. Washington, D.C.: Winston & Sons, p. 258.

Unewisse, Anne M et al. (2015). 'Observations of a Travelling Ionospheric Disturbance over'. In: *Proceedings from 15th Australian Space Research Conference*. Canberra, pp. 217–228.

Valladares, C. E. et al. (2009). 'Simultaneous observation of traveling ionospheric disturbances in the Northern and Southern Hemispheres'. In: *Annales Geophysicae* 27.4, pp. 1501–1508. DOI: 10.5194/angeo-27-1501-2009.

Van De Kamp, M., D. Pokhotelov and K. Kauristie (2014). 'TID characterised using joint effort of incoherent scatter radar and GPS'. In: *Annales Geophysicae* 32.12, pp. 1511–1532. DOI: 10.5194/angeo-32-1511-2014.

Vlasov, A. et al. (2011). 'A study of traveling ionospheric disturbances and atmospheric gravity waves using EISCAT svalbard radar IPY-data'. In: *Annales Geophysicae* 29.11, pp. 2101–2116. DOI: 10.5194/angeo-29-2101-2011.

Yang, Heng (2019). 'On the detection of ionospheric waves, relationship with earthquakes and tsunamis'. PhD thesis. Universitat Politècnica de Catalunya.

Yang, Heng, Enrique Monte-Moreno and Manuel Hernández-Pajares (2017). 'Multi-TID detection and characterization in a dense Global Navigation Satellite System receiver network'. In: *Journal of Geophysical Research: Space Physics* 122.9, pp. 9554–9575. DOI: 10.1002/2017JA023988.

Yeh, K. C. and C. H. Liu (1974). 'Acoustic-gravity waves in the upper atmosphere'. In: *Reviews of Geophysics* 12.2, pp. 193–216. DOI: 10.1029/RG012i002p00193.